

SPEECH MONOGRAPHS

Published by THE SPEECH ASSOCIATION OF AMERICA

Volume XXV

March, 1958

No. 1

| | |
|---|----|
| The Wisconsin Sequential Sampling Audience Analyzer | 11 |
| HERMAN H. BROCKHAUS and JOHN V. IRWIN | |
| Intelligibility Tests: A Review of Their Standardization, Some Experiments, and a New Test | 14 |
| JOHN WILLIAM ASHER | |
| [au] Monosyllables | 29 |
| JAMES W. ABEL | |
| The Effect of Creative Activities on the Articulation of Children with Speech Disorders | 42 |
| BARBARA M. MCINTYRE | |
| An Experimental Study of Audience Recognition of Emotional and Intellectual Appeals in Persuasion | 49 |
| RANDALL C. RUECHELLE | |
| Hermogenes on "Stock Issues" in Deliberative Speaking | 59 |
| RAY NADEAU | |
| Use of Authorities as Ethical Proof In the Talmudic Discourse | 67 |
| GERALD M. PHILLIPS | |

R 406 N 21

v. 25 1958 c. 1

SUGGESTIONS FOR CONTRIBUTORS TO SPEECH MONOGRAPHS

A portion of *SPEECH MONOGRAPHS* is devoted to the publication of articles based on original research, representing the various areas of specialization and techniques of investigation included in the field of speech.

These reports should be given in enough detail to permit the reader to know how the author carried out his investigation and how he arrived at his conclusions. Sources of evidence, conditions of observations, methods of gathering data, should be cited or described. Reports of experimental investigations should follow the usual pattern: statement of purpose, procedures used in gathering data, analysis of data, conclusions. While the inclusion of the data is essential, it should be presented concisely. Tables should be on separate pages. If statistical methods are employed, they should be named, but they need not be described if they are standard, or are adequately explained in some source to which the author can refer.

Articles in *SPEECH MONOGRAPHS* vary considerably in length. Short ones will be considered by the Editorial Board when they represent investigations the nature or scope of which permits a brief report without omitting essential details. Authors preparing reports of major studies involving extensive data and/or analysis, and anticipating that their material will exceed 8,000 words, should write to the Editor, inquiring about the availability of space and giving an estimate of the probable length of the material.

Contributors should have their manuscripts read by competent critics before submitting them for publication. When manuscripts are based on dissertations, that fact should be indicated, with the name of the director of the original research, in a footnote. In such cases, the director should review the manuscript. All copy should be proofread by at least one person besides the author before being submitted for publication. The number of words in the copy should be noted in the upper right-hand corner of the first page. Titles should be typed in full capitals, with the author's name in capitals, two spaces below. The author's affiliation, school or college (or home city, if he is not so affiliated), should appear, with normal capitalization, two spaces below his name. Footnotes should be typed on separate sheets following the last page of the manuscript. They should be double spaced, with triple spacing between items. Authors should retain a duplicate manuscript, and should send a carbon copy to the editor along with the original. The original copy should be on heavy bond paper.

All manuscripts must conform to the *MLA Style Sheet*, a copy of which can be secured from the Treasurer, Modern Language Association, 100 Washington Square East, New York 3, New York. Price: 25 cents.

Manuscripts will be read by the Editor and members of the Editorial Board before a final decision is made. Accepted manuscripts will be published in approximately the order of their acceptance. The deadline for submitting final revision of accepted manuscripts is two months before date of issue. Publication dates are March, June, August, and November. Contributors should keep the Editor informed of any changes in address.

Reprint order blanks are sent to authors; orders should be placed at once with the Arcraft Press, Columbia, Missouri. Reprints cannot be secured later.

Manuscripts should be sent to Howard Gilkinson, Editor of *SPEECH MONOGRAPHS*, 309 Folwell Hall, University of Minnesota, Minneapolis 14, Minnesota.

LLR

SPEECH ASSOCIATION *of* AMERICA

Founded, 1914
Incorporated, 1950

OFFICERS

| | |
|--|---|
| ELISE HAHN, <i>President</i> | University of California at Los Angeles |
| J. JEFFERY AUER, <i>Executive Vice-President</i> | University of Virginia |
| JOHN E. DIETRICH, <i>First Vice-President</i> | Ohio State University |
| KENNETH G. HANCE, <i>Second Vice-President</i> | Michigan State University |
| DONALD C. BRYANT, <i>Editor, The Quarterly Journal of Speech</i> | Washington University |
| HOWARD GILKINSON, <i>Editor, Speech Monographs</i> | University of Minnesota |
| KARL F. ROBINSON, <i>Editor, The Speech Teacher</i> | Northwestern University |

ADMINISTRATIVE COUNCIL

Term Expires 1958

| | |
|--|--|
| JOHN E. DIETRICH, <i>First Vice-President</i> | Ohio State University |
| KENNETH G. HANCE, <i>Second Vice-President</i> | Michigan State University |
| BARNARD HEWITT, <i>Elected by SAA, 1955</i> | University of Illinois |
| T. EARLE JOHNSON, <i>Elected by SAA, 1955</i> | University of Alabama |
| EVELYN KONIGSBERG, <i>Elected by SAA, 1955</i> | Washington Irving High School, New York City |
| THOMAS A. ROUSSE, <i>President of SAA, 1955;</i> <i>Ex Officio, 1956-1958</i> | University of Texas |
| CLARENCE T. SIMON, <i>Elected by SAA, 1955</i> | Northwestern University |
| KARL R. WALLACE, <i>Member of Finance Committee</i> | University of Illinois |

Term Expires 1959

| | |
|--|---------------------------------|
| J. JEFFERY AUER, <i>Editor, Speech Monographs, 1954-1956;</i> <i>Ex Officio, 1957-1959</i> | University of Virginia |
| ROBERT D. CLARK, <i>Elected by SAA, 1956</i> | University of Oregon |
| KENNETH G. HANCE, <i>Member of Finance Committee</i> | Michigan State University |
| WILBUR SAMUEL HOWELL, <i>Editor, The Quarterly Journal of Speech,</i> <i>1954-1956; Ex Officio, 1957-1959</i> | Princeton University |
| KARL F. ROBINSON, <i>Elected by SAA, 1956</i> | Northwestern University |
| LESTER THONSEN, <i>President of SAA, 1956;</i> <i>Ex Officio, 1957-1959</i> | College of the City of New York |

Term Expires 1960

| | |
|---|-------------------------------|
| WALDO W. BRADEN, Executive Secretary of SAA, 1954-1957; Ex Officio July 1956-June 1960 | Louisiana State University |
| THORRELL FEST, Elected by SAA, 1957 | University of Colorado |
| KENNETH G. HANCE, Executive Vice-President, 1955-1957; Ex Officio, 1958-1960 | Michigan State University |
| HENRY L. MUELLER, Editor, <i>The Speech Teacher</i> , 1955-1957; Ex Officio, 1958-1960 | University of Illinois |
| ROBERT T. OLIVER, Elected by SAA, 1957 | Pennsylvania State University |
| OWEN PETERSON, Executive Secretary of SAA, July 1957-June 1960 | Louisiana State University |
| LOREN D. REID, President of SAA, 1957; Ex Officio, 1958-1960 | University of Missouri |

Term Expires 1961

| | |
|--|---|
| WALDO W. BRADEN, Member of Finance Committee | Louisiana State University |
| ELISE HAHN, President of SAA, 1958 | University of California at Los Angeles |

Term expires 1962

| | |
|--|-------------------------|
| DONALD C. BRYANT, Editor, <i>The Quarterly Journal of Speech</i> , 1957-1959; Ex Officio, 1960-1962 | Washington University |
| HOWARD GILKINSON, Editor, <i>Speech Monographs</i> , 1957-1959; Ex Officio, 1960-1962 | University of Minnesota |

Term Expires 1963

| | |
|---|-------------------------|
| KARL F. ROBINSON, Editor, <i>The Speech Teacher</i> , 1958-1960; Ex Officio, 1961-1963 | Northwestern University |
|---|-------------------------|

SPEECH MONOGRAPHS

VOLUME XXV

MARCH, 1958

No. 1

THE WISCONSIN SEQUENTIAL SAMPLING AUDIENCE ANALYZER

HERMAN H. BROCKHAUS and JOHN V. IRWIN
University of Wisconsin

EVERY speech or performance presented to an audience elicits responses from that audience. Sometimes the responses are those which the speaker, performer, or script writer aimed to arouse; sometimes they are not. How to determine the reactions evoked by a presentation is and has been a question of concern to audience-reaction researchers. One of the first methods employed was the ballot.¹ This technique proved to be valuable in recording a general response, such as *opinion*, to a total program.

A method which provided a record of discrete reactions to an on-going program was devised in 1937 by Paul F. Lazarsfeld of Columbia University and Frank Stanton of the Columbia Broadcasting System.² These researchers developed an apparatus which recorded virtually instantaneous listener responses; they called the instrument the Program Analyzer and used it to test radio programs. Since the appearance of the original Lazarsfeld-Stanton Program Analyzer other audience-response recorders have been devised and used for various research purposes. Loran C.

Twyford's concise description of "some of the more recent audience response recorders and analyzers" includes 15 such instruments.³ Among them are the Film Analyzer used in the Instructional Film Research Program of Pennsylvania State University, and the Meier Audience-Response Recorder designed by N. C. Meier of Iowa State University.

In 1953 John V. Irwin and Herman H. Brockhaus, both staff members of the University of Wisconsin Department of Speech, developed the Wisconsin Sequential Sampling Audience Analyzer. Like other audience-response recorders, the Wisconsin Analyzer is designed to measure listener reaction to a given program. There are several features which characterize this analyzer and the method it employs; some are similar to those of other recorders and analyzers while other features are unique:

First, the Wisconsin Analyzer records the conscious responses of the subjects. Like most response recorders the Wisconsin Analyzer relies upon subjective evaluation on the part of the listeners whose responses are being measured.

Second, the Analyzer provides for measurement of listener reaction on either a three-point or a two-point scale. For example, on a three-point scale, "plus," "neutral," "minus"

¹ William A. D. Millson, "Experimental Work in Audience Reaction," *Quarterly Journal of Speech*, XVIII (1932), 13-30.

² Tore Hallonquist, "'Big Annie' Goes to Work," *Advertising and Selling*, XXXVIII (August, 1945), 69, 156, 158, 160.

³ Loran C. Twyford, "Profile Techniques for Program Analysis," *Audio Visual Communication Review*, II (1954), 243-262.

could be employed; on a two-point scale, "involvement" and "non-involvement."⁴

Third, the Wisconsin Analyzer uses a sampling method. The subjects indicate their responses continuously throughout the program, not just on given signals; the Analyzer takes a *sample* of each subject's response at regular intervals without the subject's being aware of the precise moment his reaction is sampled. The samples are taken in sequence: Subject 1, Subject 2, Subject 3, etc.

Fourth, the Analyzer is designed to provide two separate records: (1) a record of group responses, and (2) a record of discrete individual responses. Either or both records may be obtained from the Analyzer.

Fifth, the Wisconsin Analyzer can furnish two kinds of data: (1) a cumulative record of responses for the entire program, and (2) a short-term record of responses for any segment of the program. The two kinds of response data can be obtained for either the group, for individual subjects, or for both.

In describing more fully the properties and functions of the Wisconsin Sequential Sampling Audience Analyzer it will be helpful to relate portions of the explanation to an actual program in which the Analyzer was used. The project selected for reference was a demonstration-study conducted for the Wisconsin Telephone Company, on the University of Wisconsin campus in Madison in April, 1955. In this demonstration-study a Telephone Company public relations speaker presented a Telephone Company public relations speech to three separate University audiences.

The explanation of the operational components of the Analyzer and their use can conveniently be divided into three parts: (1) the apparatus used by the listeners to indicate their responses, (2) the mechanism used to sample the listeners' responses, and (3) the instru-

ments employed to record the sampled responses.

The equipment used by the listeners consists of 24 identical response boxes. Each member of the audience who is a subject is given one of the boxes (Figure 1) which he holds in his hands or rests on the arm of his chair. The response box, made of aluminum, is 3 x 4 x 5 inches in size. Mounted on the upper face of the box is a double-pole, three-position lever switch which is spring-loaded so that normally the lever is upright. In order to move the lever back or forth from the normal mid position to either of the other two switch positions, slight pressure must be applied.

A label showing the measurement scale employed is pasted on the box, adjacent to the switch lever. In the Telephone Company demonstration-study the aim was to measure *interest* in the program; the label read: "Degrees of Interest—HIGH, MEDIUM, LOW."

Before the start of a program to be studied the positions on the scale are defined for the subjects as accurately as possible. In the Telephone Company project a "high" degree of interest was defined as "giving attention without any effort." A "low" degree of interest was described as "having to work to pay attention." For want of a better definition a "medium" degree of interest was simply described as "halfway between giving attention without effort, and working to pay attention."

Besides definitions of the positions on the scale, careful instructions are given the listeners who will operate the response boxes. The subjects are informed that throughout the program they are to move the lever from one position to another *at will*; they are to move the lever any time they become aware of a change in their response.

⁴ University of Wisconsin Television Laboratory Research Bulletin No. 6, *In Rehearsal: Twelfth Night—Program Evaluation*, September 1, 1956, Madison, Wisconsin.



FIGURE 1. A response box used by the subject to indicate his responses. The normal position of the lever is at the mid position, as shown.

After these instructions the subjects are given an opportunity to ask questions and also to practice manipulating the levers.

Turning from an exposition of the response boxes to an explanation of the mechanism used to sample the listeners' responses, it might be pointed out again that although the subjects indicate their responses continuously throughout the program, the Audience Analyzer merely *samples* each subject's reaction periodically. At regular intervals—once every five seconds when the fastest motor speed is used—the Analyzer takes a given subject's response at the moment and feeds it to the recording mechanism. Each box operator is sampled periodically in this manner.

The sampling, as stated previously,

is done in sequence. First the response of Subject No. 1 is sampled, then that of Subject No. 2, then Subject No. 3—and so on through the sequence of box operators. At the end of the first sampling cycle, five seconds in duration (or longer if a slower motor speed be used), a time mark is registered. Then a second sampling cycle begins immediately. The reaction of Subject No. 1 is sampled again, then No. 2, etc. This sampling process is repeated over and over again from the beginning to the end of the program, interrupted only momentarily by the recording of time-marks which separate the sampling cycles.

The equipment which accomplishes this periodic sampling may be called the Analyzer proper (Figure 2). Its

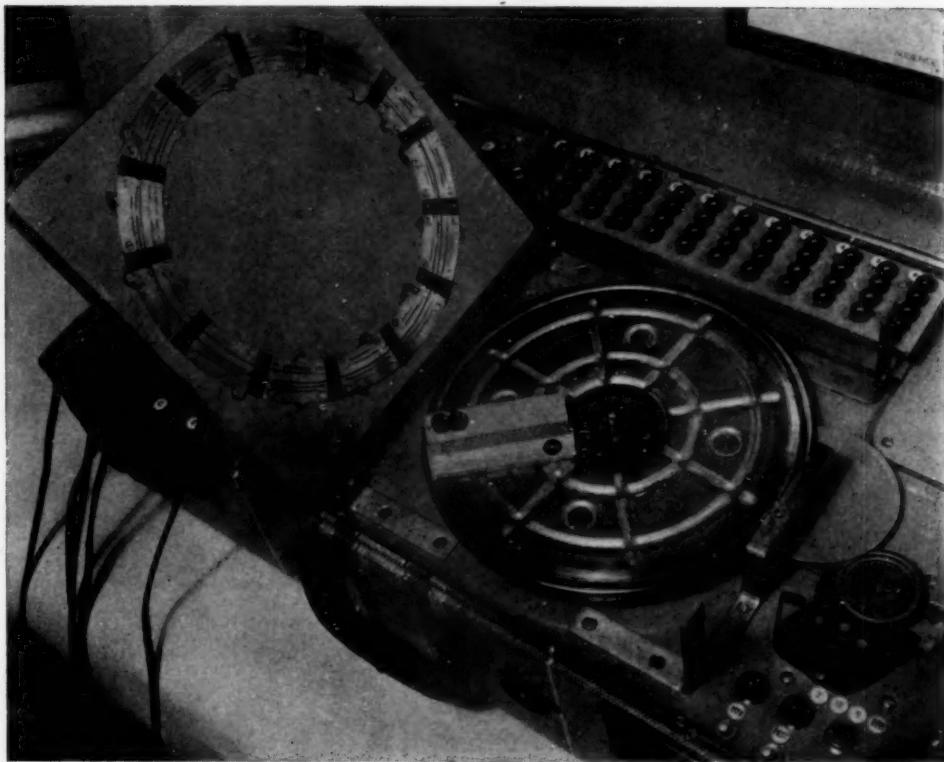


FIGURE 2. The Analyzer proper, the mechanism which samples the listeners' responses sequentially and feeds them to the recording instruments. The wooden platform has been inverted to show the commutator plates; normally the platform covers the turntable, plates down.

principal components are a wooden platform on whose under side are two separate sets of copper commutator plates mounted in concentric circles, a phonograph turntable with two carbon brushes, and a motor plus step-down drive.

Each circle of commutator plates is comprised of 13 segments with plastic dividers between them for insulation. Twelve of the 13 copper plates in each circle are connected with response boxes. A three-conductor cable is used to connect each plate with its corresponding box. The 13th plate, located between the 12th and the 1st and designated as "O," is connected with an indicator which registers the number of sampling cycles.

The outer circle of copper plates is

for one circuit of 12 boxes, which was arbitrarily designated as the "red" circuit in the Telephone Company demonstration-study, while the inner circle is for an identical but separate second circuit of 12 boxes, designated the "white" circuit. This dual-circuit design makes possible the separate sampling of discrete responses from two groups of 12 persons each.

Besides the two commutator circles, the sampling mechanism includes a pick-up rotor in the form of a phonograph turntable on which has been mounted two carbon brushes, one brush for each circle. The turntable is driven by a conventional phonograph motor whose three speeds have been reduced by a custom-designed friction drive. The resultant turntable, or rotor, speeds

are 12, 7, and 5 rpm; their respective sampling cycles are 5 seconds, 8.6 seconds, and 12 seconds approximately. In the Telephone Company demonstration-study the rotor was operated at the maximum speed; thus, as indicated previously, the response of each subject was sampled once every 5 seconds.

An explanation of the Analyzer in operation will illustrate more clearly how the sequential sampling is accomplished. With the wooden platform in its proper position, face down, the turntable revolves. As it revolves, the two separate brushes make sequential contact with each plate in their respective circuits. When a brush contacts a given plate it makes possible an electrical connection between the corresponding response box and one of two single-throw relays.

If the response box lever be in the "high" position, the "high" relay will close and complete an electrical connection with a recording instrument. If the response box lever be in the "low" position, the "low" relay will close and complete a different connection with a recording instrument. If the response box lever be in the "medium" position, the circuit remains open and a "medium" response is indicated. Since there are two separate circuits, "red" and "white," a total of four relays is required. These relays are 6-volt A.C., double-pole, single-throw in type and are mounted in a normally open position. A schematic diagram of the Analyzer is shown in Figure 3.

The relays establish connections between the Analyzer proper—the mechanism used to sample the listeners' re-

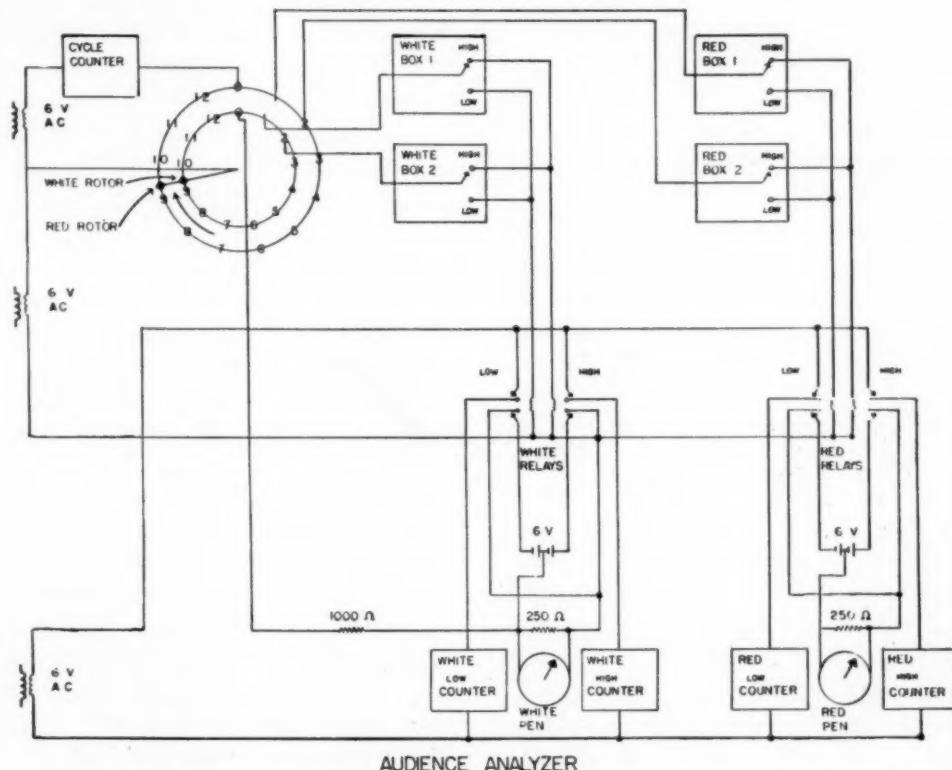


FIGURE 3. A schematic diagram of the Wisconsin Sequential Sampling Audience Analyzer.

sponses—and the instruments employed to record the sampled responses. These instruments are: (1) four electric counters, and (2) a double-channel recording meter or ink-writer. One pole of each of the four relays is connected with a corresponding electric counter which totals the number of times the connected relay is actuated. The second pole of each relay is connected with appropriate terminals of pens in the ink-writer. The counters give a record of the group responses sampled while the ink-writer provides a record of the individual responses sampled.

A more detailed exposition of the recording instruments and their function will help to explain the kind of records they provide. The four electric counters are of the four-digit Veeder-Root resettable variety. They afford a cumulative group response record by tallying the number of "high" and "low" samples for each of the two circuits. In addition, a fifth electric counter, connected with the "O" plate in the "red" circuit, registers the number of sampling cycles by recording the turn-table, or rotor, revolutions.

Information provided by this fifth counter makes it possible to compute the total number of "medium" samples. Since 12 samples per circuit are taken in each sampling cycle, the product of 12 times the number of cycles is the total number of samples per circuit taken during a given program. By subtracting from this total the sum of the "high" and "low" samples for a circuit, the number of "medium" samples for that circuit can be obtained.

In the Telephone Company demonstration-study, four observers took readings of the electric counters every 15 seconds, copying down the cumulative totals of the "high" and "low" samples. This written record made possible an

easy computation of response variations during the speech within quarter-minute intervals. The 15-second time interval was selected on the basis of preliminary investigation and the experience of other workers as being short enough to reveal significant changes.

As stated previously, the electric counters provided a group response record in the Telephone Company project. A record of discrete individual response samplings was also registered. This was accomplished by using a Brush BL-202 Double-Channel Oscillograph. Basically this recording meter consists of a paper-feeding mechanism and two pens each of which is capable of moving in a restricted arc. Chart paper is pulled from a roll at a constant speed under the two wet-ink capillary-feed pens. Three chart speeds are available: 5, 25, and 125 millimeters per second. In the demonstration-study the 5-millimeter speed was used. The chart paper, incidentally, is printed to show two separate channels, one for each pen.

The pen for one channel records the responses sampled from the 12 boxes in the "red" circuit while the pen for the other channel records the responses sampled from the 12 boxes in the "white" circuit. When the trace made by a pen is on the base line, a "medium" response is recorded; when the trace deflects above the base line, a "high" response is registered; and when the deflection is below the base line, a "low" response is indicated.

The trace for one channel of the ink-writer, the "white" circuit channel, is divided into segments by blips, or short vertical lines. These blips mark the beginning and the end of each sampling cycle and are drawn when the rotor brush for the "white" circuit makes momentary contact with the "O" plate of that circuit. This contact com-

pletes causes an inst which
Exce writer motor, tery. in par resisto ing, th markin chart

As
duced
ord
pled
subje
overla
it is
spons
struct
samp

To
ence
recor
match
plish
recor
react
trans
back
possi
time
in th

In
for
of
tatio
scrip
inte
each
recor
ed
down
cou

T

pletes a 6-volt A.C. connection which causes the pen to fluctuate rapidly for an instant, drawing a short heavy mark which intersects the base line.

Except as noted above, each ink-writer pen is driven by a magnetic pen motor, powered by a $7\frac{1}{2}$ -volt "C" battery. A 250-ohm resistor is connected in parallel with each pen motor. These resistors provide for an optimum damping, thereby squaring off the deflection markings on the graph, making the chart easier to read.

As mentioned earlier, the chart produced by the ink-writer provides a record of each individual response sampled during the entire program, for 24 subjects or box operators. By using an overlay in conjunction with the chart it is possible to read the discrete responses of any one subject and construct a separate record of that subject's sampled responses.

To produce a precise analysis of audience reaction to a given program, the record of responses sampled must be matched with the program. To accomplish this end the program is sound-recorded at the same time the audience reactions are sampled. Later a typed transcription is made. Timing a playback of the sound-recording makes it possible to indicate on the manuscript time intervals which correspond to those in the record of sampled responses.

In the Telephone Company project, for example, a tape-recording was made of each of the three live speech presentations; from these recordings manuscript transcriptions were typed. Time intervals of 15 seconds were marked on each manuscript to correspond to the record of the electric counters, recorded by the four observers who copied down readings of the "high" and "low" counters every quarter minute.

The data collected in the Telephone

Company programs were not given statistical treatment since the project was regarded simply as a demonstration of the Wisconsin Analyzer as an interest-measuring instrument and was not thought of as experimental research. However the data were organized in a systematic fashion and were reported in ways that related both the group response record and the individual response record to the speech.

The first method of reporting applies to the response record for the entire group. This record is matched with the program by indicating in the margin of the typewritten transcription the percentage of "high," "medium," and "low" degree-of-interest responses sampled during each 15-second interval of the speech. Figure 4 is a facsimile of the first page of the speech delivered to the 3:30 p.m. audience in the demonstration-study, plus response data. The numbers in parentheses in the body of the manuscript are the 15-second markers; these correspond to the points at which the counters were read by the observers.

The sets of percentage figures in the margin are identified by numbers in parentheses which correspond to those in the manuscript. Number 1 in the manuscript marks the first 15-second segment of the speech beginning with the first word in the talk, "Well," and ending with the word "you." According to the marginal record, identified by (1), during the first 15 seconds of the speech $36\frac{1}{4}\%$ of the group responses sampled were "high"; $57\frac{1}{2}\%$ were "medium"; and $6\frac{1}{4}\%$ were "low." The method of reporting illustrated on this page was used throughout the 3:30 p.m. speech transcription.

The second method of reporting, as illustrated in Figure 5, also shows a group response. It differs from the first

A NATION AT YOUR FINGERTIPS

William Peterburs

April 27, 1955 -- Madison, Wisconsin

3:30 Audience

Well this is rather interesting, being a part of an experiment. It's a little nerve-wracking; so if I appear to be a little bit nervous, it's because I actually am. It's the first time I've ever been a guinea pig, shall we say.

(1) Well, today I'd like to talk to you about your telephone service. I know it's a very important part of your everyday life, and that it has been for at least the last five or six years, ever since you were a freshman in high school, probably. And from the reports that I get from (2) Charlie Goulet here in Madison I know the University students have a tremendous use of the telephone.

Now if we go back to the early days and try and visualize a telephone system of the -- those days, every telephone in the system required a pair of wires connected to it from every other telephone in the system. Let's take a six-party telephone system, for there are only six people in the town and each one of them has a telephone.

(3) (4) There would be five wires connected to their -- five pair of wires connected to their house, one from every other telephone in the system.

It works very well, but those six telephones would require 15 pair of

FIGURE 4. The first page of a speech manuscript plus group data, illustrating a method of reporting the response record for the entire group of subjects. The marginal numbers indicate the percentage of "high," "medium," and "low" degree-of-interest responses sampled during four separate consecutive 15-second intervals, indicated by the numerals in parentheses.

method in that it presents a separate response record for each half of the subjects in the study. In the Telephone Company project the speech given at 1:20 p.m. was delivered to an audience of 24 persons comprised of 12 men engineering students and 12 women students, or co-eds. The men were assigned response boxes on the "red" circuit, while the women were given boxes

on the "white" circuit. Since the counters gave separate readings for each circuit, it was possible to present a separate report of the engineers' responses and of the co-eds' responses. By placing the percentage figures for the two halves of the audience side by side in the margin of the speech transcription, a comparison of the responses of the two sub-groups is facilitated. A dif-

A NATION AT YOUR FINGERTIPS

William Peterburgs

April 27, 1955 -- Madison, Wisconsin

1:20 Audience

I don't know about this being an experiment today. If I appear a little bit nervous it's because I actually am. This is the first time I've been a guinea pig, shall we say, for an experiment of this type.

(1) Today I'd like to talk to you about telephone service. We know it's an important part of all of your everyday lives. We know too, from a report I've had from Charlie Goulet, University students' use of the telephone is quite amazing as to the amount of use that is made of that instrument.

Eng. Co-ed

(1)

| | | |
|---|------|------|
| H | 3.5 | 10.0 |
| M | 95.0 | 82.5 |
| L | 2.5 | 7.5 |

| | | |
|-----|------|------|
| (2) | 30.0 | 40.0 |
| | 70.0 | 52.5 |
| | 0.0 | 7.5 |

Now talking about telephone service can get quite complicated in a hurry. In fact talking about telephone service for six telephones, as (3) this drawing illustrates, appears quite complicated. And this drawing itself and the network of wires that I have connected are a replica (4) of what a telephone system of about 1876 or '77 looked like. The system worked very well, but notice that every telephone in the system

| | | |
|-----|------|------|
| (3) | 40.0 | 57.5 |
| | 60.0 | 32.5 |
| | 0.0 | 10.0 |
| (4) | 57.5 | 52.5 |
| | 35.0 | 42.5 |
| | 7.5 | 5.0 |

had to have a pair of wires connected to it from every other telephone in the system. As I said, that system worked very well, but those six (5) -- well six telephones took 15 pair of wires to interconnect them. If we had -- still had this type of interconnection in existence today,

| | | |
|-----|------|------|
| (5) | 75.0 | 52.5 |
| | 22.5 | 42.5 |
| | 2.5 | 5.0 |

the number of telephones that we have in Wisconsin alone, we would be

FIGURE 5. The first page of a speech manuscript plus group data, illustrating a method of reporting the separate response records for each half of the subjects. The audience for this speech was 12 engineering students and 12 women students, or co-eds.

ference between the responses of the engineers and the responses of the co-eds in the fifth 15-second segment of the talk is readily apparent.

The third method of reporting data, shown in Figure 6, differs from the previous two in that it presents *individual* response samplings. The group record is also revealed so that the reactions of the individual may be compared with those of the group. Both records—in-

dividual, at 5-second intervals; and group, at 15-second intervals—are placed in the margin of a typewritten transcription of the program.

The individual response record, as indicated earlier, is obtained from the recording meter chart. In the Telephone Company demonstration-study an individual record was compiled for each of eight selected subjects. The method of reporting the individual re-

A NATION AT YOUR FINGER TIPS

William Peterburg

April 27, 1955 -- Madison, Wisconsin

1:20 Audience

I don't know about this being an experiment today.

If I appear a little bit nervous it's because I actually am. This is the first time I've been a guinea pig, shall we say, for an experiment of this type.

(1)

Today I'd like to talk to you about telephone service. We know it's an important part of all of your everyday lives. We know too, from a report I've had from Charlie Goulet, University students' use of the telephone (2) is quite amazing as to the amount of use that is made of that instrument.

Now talking about telephone service can get quite complicated in a hurry. In fact talking about telephone service for six telephones, as this drawing illustrates, appears quite complicated. And this drawing itself and the network of wires that I have connected are a replica of what a telephone system of about 1876 or '77 looked like. The system worked very well, but notice that every

Subject: Co-ed No. 2
Course of Study: General
Course (B.A. Degree)
Classification: Sophomore

| Individual's Level of Interest | % for Group |
|--------------------------------------|-------------------|
| M | |
| M | |
| M | |
| M | (1) |
| H | H 6.25 |
| M | M 88.75 |
| M | L 5.00 |
| M | |
| H | (2) |
| H | 35.00 |
| H | 61.25 |
| H | 3.75 |
| H | |
| H | |
| H | |
| H | L |
| H | (3) |
| H | 48.75 |
| H | 46.25 |
| H | 5.00 |
| H | |
| H | |
| H | |
| H | (4) |
| H | 55.00 |
| H | 38.75 |
| H | 6.25 |

FIGURE 6. The first page of a speech manuscript plus individual and group data, illustrating a method of reporting the response record for an individual, at 5-second intervals, as well as the response record for the entire group of subjects, at 15-second intervals.

sponses, as sampled every 5 seconds, is illustrated in Figure 6, which shows the first page of the report for Co-ed No. 2. The subject's responses are recorded in three parallel columns, one column for each position on the level-of-interest scale. The "H" in the first column indicates a "high" sample; the "M" in the second column, a "medium" sample; the "L" in the third column,

a "low" sample. The vertical location of the letters in the columns, in relationship to the typewritten lines in the transcription, shows what portion of the program elicited a particular response. The percentage figures in an adjoining column show the group response for 15-second intervals; this record was obtained from the counter readings.

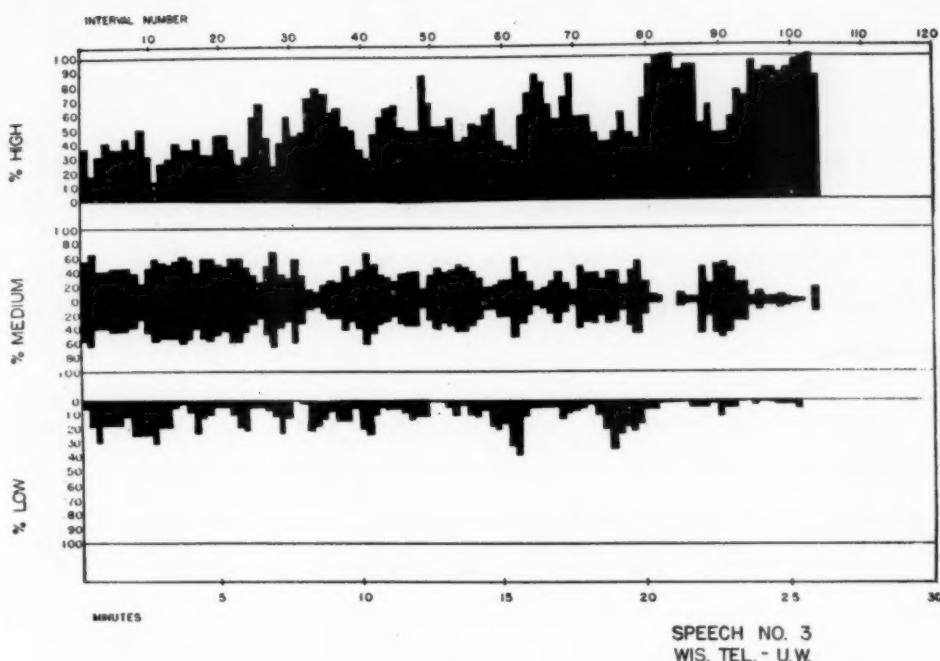


FIGURE 7. A profile of group responses, illustrating a method of reporting the response record for all subjects for an entire speech. Each bar represents the response sampled during a 15-second interval.

The fourth and final method of reporting data collected with the Wisconsin Analyzer, illustrated in Figure 7, is another group report. It is a profile which shows the percentages of "high," "medium," and "low" samples registered throughout the program by all of the subjects. Figure 7 is a profile record of responses to the 3:30 p.m. speech presentation in the Telephone Company project. Three parallel divisions comprise the profile: "high," "medium," and "low." In the "high" portion the bars extend above the base line; in the "low" portion they extend below the base line; while in the "medium" portion they straddle the base line. This profile design reveals at a glance both the direction and the amount of change in interest on the part of the group.

Each bar of the profile represents the responses sampled during a 15-second

interval. The "interval numbers" on the graph help to identify the quarter-minute periods. The first bar indicates the 15-second interval starting with the first word in the speech, to the moment the counters were first read by the observers 15 seconds later. By matching the bars on the profile with the corresponding numbers in the body of the typewritten transcription (Figure 4) it is possible to relate any portion of the graph to the corresponding portion of the speech. For example, an examination of the profile shown in Figure 7 will reveal one 15-second interval in the conclusion of the speech, interval "103," in which the interest response was 100% "high." An inspection of Figure 8, which is a facsimile of the last page of the 3:30 p.m. speech transcription and report, will show what portion of the speech obtained

| | |
|--|---------|
| (99) | (99) |
| out our ever touching the steering wheel or stepping on the gas, this | H 87.50 |
| transistorized brain will take us in our automobile to our destination, | M 8.75 |
| merely by following transistorized circuits that have been laid in the | L 3.75 |
| highway, without ever having an accident or without ever exceeding the | |
| speed law. | |
| (100) | (100) |
| And then in the ultimate, as Dr. Osborne just said -- Dr. Osborne | 90.00 |
| was formerly the chief engineer for the American Telephone and Telegraph | 5.00 |
| Company -- he has said that whenever a human being, ultimately, whenever | 5.00 |
| (101) | (101) |
| a -- whenever a human being is born any place in the world, he will be | 96.25 |
| assigned a telephone number, and that will be his telephone number for | 3.75 |
| life. And then when he's old enough to talk, he'll be given a little | 0.00 |
| gadget that has on one side of it 10 little buttons, and on the other side | |
| (102) | (102) |
| a little screen. Whenever he wants to call anybody, he'll take this | 98.75 |
| gadget out of his pocket, punch out his friend's telephone number on these | 1.25 |
| 10 little buttons, turn the gadget over, and he'll be able to hear his | 0.00 |
| friend and see him in both color and in three dimensions. (Laughter) And | |
| (103) | (103) |
| then, if he can't hear or see his friend, he will know that he is | 100.00 |
| dead, (Laughter) because that's the only way it would work -- and that's | 0.00 |
| how I am right now -- dead. And so there you are. | *(104) |
| | 85.00 |
| | 15.00 |
| | 0.00 |

*This final reading of the counters was made after the speech had been completed.

(3:30 Audience)

FIGURE 8. The last page of a speech manuscript plus group response data. A comparison of the response record for all subjects on this page with the last five bars of the profile in Figure 7 illustrates how the speech and the profile may be matched.

this unanimous "high" response. The drop from a "high" of 100% in interval "103" to 85% in interval "104" probably can be explained by the fact that the final counter reading was made several seconds after the last word had been spoken. Undoubtedly a number of subjects moved the levers of their response boxes to the normal upright

position as soon as the speaker had finished talking.

For those persons who are interested not only in the design and operation of equipment but also in its cost, the following table gives a cost estimate of the Wisconsin Sequential Sampling Audience Analyzer. Figures include labor costs where applicable.

| Item | Number | Cost | Total |
|-------------------------------|----------|------------------|-----------------|
| Response box including switch | 24 | \$ 2.50 each | \$ 60.00 |
| Cable (3-conductor) | 2400 ft. | 3.00 per 100 ft. | 72.00 |
| Sampling mechanism | 1 | 290.00 | 290.00 |
| Electric counters | 5 | 12.00 each | 60.00 |
| Recording meter | 1 | 800.00 | 800.00 |
| | | | <hr/> \$1282.00 |

INTELLIGIBILITY TESTS: A REVIEW OF THEIR STANDARDIZATION, SOME EXPERIMENTS, AND A NEW TEST*

JOHN WILLIAM ASHER
U. S. Office of Education

IN an experiment concerning the phenomenon of intelligibility, the results can be no more accurate than the degree of accuracy with which the intelligibility lists were first equated. If more precise methods of equating intelligibility lists can be established, then subsequent experimentation in voice communications will become more efficient.

The purpose of this investigation was to study the previous methods of equating multiple choice intelligibility lists, to develop a method of constructing more precisely equated listening intelligibility tests, and to develop a new type intelligibility test.

REVIEW OF PREVIOUS EXPERIMENTS

Statistical Considerations. Several reports and articles^{1,2,3} have summarized the history and results of previous ex-

perimentation in intelligibility testing. This article will not attempt to repeat this work. None, however, in the writer's opinion, has reviewed adequately, especially in light of recent advances in statistical technique,^{4,5} a fundamental problem of any testing procedure, that of equating the means of alternate forms of a test. Two other methodological problems, the use of per cent scores as criterion measures and the round robin procedure, will also be reviewed.

One of the first attempts to determine empirically the equivalence of the means of a set of alternate forms of an intelligibility test was done by the Harvard Psycho-Acoustic Laboratory in an experiment⁶ with E-lists (write down lists) which had been constructed in that laboratory. Ten of the lists were arbitrarily paired, and each talker tested for intelligibility read one pair. Since each of the five pairs was read by a number of talkers, it was possible to compute the difference between the mean scores of the lists within each pair. Multiple *t* tests were used to compare the means. No significant *t* ratio was found between any pair of lists. It should be noted that the means of all the lists were not compared for equivalence as would have been done in an

*This study was performed in partial fulfillment of the requirements for the Ph.D. degree at Purdue University. The research was carried out under contract with the Office of Naval Research, U. S. Naval Training Device Center, as Contract N6ori-104, T.O. II, Project 20-F-8, of which this experiment constituted Technical Report No. SpecDevCen 104-2-44. Professor E. R. Ryden was chairman of the writer's graduate committee and Professors M. D. Steer and T. D. Hanley were contract project directors.

¹ John W. Black, "Accompaniments of Word Intelligibility," *J. Speech and Hearing Disorders*, 17 (1952), 409.

² John W. Black, *et al.*, "Applications of Multiple-Choice Speech Intelligibility Tests in the Evaluation and Use of Voice Communication Equipment," U. S. School of Aviation Medicine, Joint Report NM 001 064.01, Report No. 19, Pensacola, Florida, 1953.

³ T. D. Hanley and M. D. Steer, "Intelligibility of Voice Communication," U. S. Naval Training Device Center Project 20-F-8, Purdue Research Foundation, Technical Report No. SpecDevCen 104-2-51 (1956).

⁴ W. G. Cochran and G. M. Cox, *Experimental Designs* (New York, 1953).

⁵ A. L. Edwards, *Experimental Design in Psychological Research* (New York, 1950).

⁶ M. H. Abrams, *et al.*, "The Use of Words and Sentences in Testing the Intelligibility of Telephone Talkers," OSRD No. 4023, IC-69 (1944).

analysis of variance. Also, neither the sums of squares due to differences among talkers nor the sums of squares due to talkers \times lists' interaction was removed from the error term used to test the differences among the means. Nor were the sums of squares due to listeners or any of the listener interactions removed from the error term. However, the techniques to accomplish this were not readily available at the time these studies were done.

In an attempt to evaluate the equivalence of the means of all the lists together, each of three experienced talkers read lists one through ten to seven listeners.

An analysis of these data showed that the variance attributable to lists was not significantly larger than the "remainder" variance due to experimental error (the variability of the talkers), so that this experiment, like the others, gives excellent grounds for believing the E-lists to be equivalent in difficulty.⁷

This is a unique application of the analysis of variance, the main effect of lists being compared to the main effect of talkers. Since it is known that talkers vary considerably among themselves^{8,9} with respect to intelligibility, if the variance attributable to lists is even larger though not significantly larger, it seems quite likely that this experiment actually demonstrates the non-equivalence of the means of the lists.

Two other experiments reported by the Harvard Psycho-Acoustic Laboratory purportedly give evidence of the equality of the means of sentence intelligibility lists. In the first¹⁰ a number of talkers each read two lists consisting of 20 sentences apiece, uniformly con-

structed from a large number of items. "The correlation between the individual scores on the two readings was .85."¹¹ It is erroneous to conclude that ". . . this may be considered satisfactory evidence for the essential equivalence of the sentence lists."¹² The highest score of one set of data need not even overlap with the lowest score of the other set of data, or vice versa, to produce such a correlation.

The second report¹³ tells of an experiment in which 24 talkers each read two lists of 20 sentences. The correlation between the individual scores on these two readings was .82. Again, the conclusion that ". . . a correlation of this size may be considered good evidence for the approximate equivalence of the sentence list,"¹⁴ is erroneous for the same reason as given above.

The Waco Voice Communication Laboratory^{15,16,17} constructed 22 lists of 24 words each with approximately equal means and standard deviations, the word difficulties having been determined by previous experimentation. These lists were tested on about 800 speakers and listeners in three groups; and with the statistical methods used, no differences were found among the lists, and, therefore, the lists were considered to be equivalent.

The 24 item word lists were later shortened to 24 12-word lists. The 24-word lists and 12-word lists were both

⁷ *Ibid.*, p. 3.

⁸ *Ibid.*, p. 4.

⁹ M. H. Abrams, *et al.*, "The Use of Words and Sentences in Testing the Intelligibility of Telephone Talkers," OSRD No. 4023, IC 69 (1944).

¹⁰ *Ibid.*, p. 7.

¹¹ C. H. Haagen, "Intelligibility Measurement: Techniques and Procedures Used by the Voice Communication Laboratory," OSRD No. 3748 (1944).

¹² C. H. Haagen, "Intelligibility Measurement: Twelve-Word Tests," OSRD No. 5414 (1945).

¹³ C. H. Haagen, "Intelligibility Measurement: Twenty-Four Word Multiple-Choice Tests," OSRD No. 5567 (1945).

¹⁴ T. D. Hanley and M. D. Steer, *op. cit.*

¹⁵ M. H. Abrams, *et al.*, "Factors Related to the Intelligibility of Talkers in Noise," OSRD No. 4023, IC-60 (1944).

write-down tests. Haagen further revised the above tests and constructed the 24-word multiple choice test and a 12-word multiple choice test. These multiple choice tests are asserted to measure intelligibility as effectively as the write-down test. Intelligibility scores of 72 speakers as measured by a write-down test and a multiple choice test (same words) correlated .72.

More recently, Black¹⁸ has constructed two more sets of multiple choice intelligibility tests to supplement the Haagen Waco Tests. They are called Forms C and D to follow the Haagen nomenclature of Forms A and B. The Black (Pensacola) tests were constructed in much the same manner as the Haagen (Waco) tests.

Black stated that "The trial lists were equated with respect to the mean level of difficulty of the items (not variance) . . ." The lists were again re-constituted on the basis of the multiple choice values of the items, which were equated for mean difficulty. The revised multiple choice intelligibility tests were read by 12 panels of Naval R.O.T.C. students, 12 per panel, in noise (round robin procedure). Listeners and speakers were subjected to 110 db of simulated aircraft noise. The speaker lists and speaker stations were rotated in the administrations of the test to the 12 panels. Thus, dissimilarity among head sets, station boxes, or acoustic conditions at the listening stations was equalized among the speaker lists. These panels also read and listened to the Harvard phonetically balanced tests and Haagen's Form A test during this experiment. An analysis of variance was made of the experiment with Black's Form C and D lists. The same

analysis was made on the Form A and the phonetically balanced tests.

According to Black's analyses the mean square attributable to lists was not significant for Form C (quiet or noise), Form D, or the phonetically balanced lists. The *F* ration was significant at the five per cent point in the analysis of the lists for the Form A test. These analyses are highly conservative for reasons which will be discussed later. Since the purpose of the experiment was, if possible, to reject the null hypothesis of equality among the means of the lists within the various forms, the most powerful test available should have been used.

Haagen, a principal investigator at the Waco Voice Communication Laboratory, in discussing the methods used to determine equivalence among the means of the Waco Intelligibility Tests says:

Two major sources of variability contribute to the observed differences between tests, the difficulty of words that make up the several tests and the ability of the speakers who speak the tests. Since each speaker reads only one test it is not possible to obtain independent estimates of these two variables. This, however, is of no concern if the obtained values of *F* are significant. A more precise analysis would contribute only a lower confidence level at which the hypothesis of unequivalent [sic] tests could be rejected. No significant differences between test means were observed.¹⁹

This is an inexact statement; listener variability is an important source of variance in intelligibility testing.²⁰ It is of major concern if a precise analysis cannot be made when the *F* test of the

¹⁸ C. H. Haagen, "Intelligibility Measurement: Twenty-Four Word Multiple-Choice Tests," *op. cit.*

¹⁹ This observation can be verified in any intelligibility testing experiment involving several randomly selected talkers. One such experiment from the Purdue laboratory is: Irwin Miller, Theodore Cloonan, and John W. Asher, "Speech Intelligibility Characteristics of Army Respirators," Chemical Corps Contract DA 18-108-CML-4037, Report No. 7 (1954).

²⁰ John W. Black, "Multiple-Choice Intelligibility Tests," U. S. School of Aviation Medicine, Joint Report NM 001 064.01, Report No. 17 (1953).

hypothesis of equality of the means cannot be rejected. With an improper analysis or experiment such as this, the mean square used for error is larger than the true error would be. If differences are found using these error terms, it is likely that true differences do exist. However, if, as in this case, the hypothesis of no differences cannot be rejected, the experiment gives no information as to whether any true differences exist among the means. Fortunately, experimental designs are now available²¹ which do not inextricably mix up, or confound, speaker and listener effects.

Throughout all of the above literature on intelligibility testing the criterion used in the various analyses seems to have been the per cent correct or incorrect. This follows Fletcher's original work²² in which the articulation score was expressed in per cent. However, Fletcher cautioned against the use of experiments in which the per cent score means were less than 20 per cent or more than 80 per cent. Statisticians, for instance Edwards,²³ have cautioned against the use of percentage data in analyses because of the tendency for the means to be proportional to the variances, causing heterogeneity of variance. This phenomenon is readily observable in experiments performed by Black²⁴ in which, as the average intelligibility scores approach 100 per cent, the standard error of the mean decreases rapidly.

Methodological Considerations. The round-robin procedure widely used in communication training programs, is an ideal way to train speakers and listeners, but as an experimental design it is essentially unanalyzable in any precise

manner, since listener groups continually change. The analysis is not a complete "repeated measurements" design, nor a complete "independent group" design. Nonetheless, intelligibility research has been reported from several laboratories (including the laboratory in which the current experiment was performed) in which this inexact analysis was employed. A typical analysis is used by Black²⁵ to analyze an experiment in which there were twelve speakers, 144 list-speaker combinations. While a given talker was speaking, the other eleven members of the panel listened and marked intelligibility tests. Thus, there are three main effects in the experiment: speakers, with 11 degrees of freedom (df); listeners, with 10 df; and lists, with 11 df. There would be 1584 observations in all. However, the eleven listeners are never the same from talker to talker, yet the criterion scores are highly correlated. An accurate analysis is virtually impossible. The solution given by Black is typical. The listener scores for any one talker-test combination are pooled, giving 144 scores. There is a main effect of "lists" with 11 df, a "remainder" with 132 df, and a total of 143 df. Analyses done at the Purdue Voice Communication Laboratory using designs which test all main effects and interactions^{26,27,28} have shown that this "re-

²⁵ John W. Black, "Multiple-Choice Intelligibility Tests," *op. cit.*

²⁶ John W. Asher, Theodore Cloonan, and Irwin Miller, "Speech Intelligibility Characteristics of Army Respirators," Chemical Corps contract DA 18-108-CML-4037, Report No. 9 (1954).

²⁷ John W. Asher and T. D. Hanley, "An Experiment Concerned With Listener Accuracy in an Environment of Relevant Conflicting Voice Signals," U. S. Naval Training Device Center Project 20-F-8, Purdue Research Foundation, Technical Report No. 104-2-39 (1954).

²⁸ Theodore Cloonan, Robert C. Bilger, and John W. Asher, "A Comparison of Three Listening Environments," Chemical Corps Contract DA 18-108-CML-4037, Report No. 8 (1954).

²¹ A. L. Edwards, *op. cit.*

²² Harvey Fletcher, *Speech and Hearing* (New York, 1929).

²³ A. L. Edwards, p. 199.

²⁴ John W. Black, "Applications of Multiple-Choice Intelligibility Tests," *op. cit.*

mainder" term will contain many effects the mean squares of which are much larger than those of the true error terms. Thus, any "remainder" term obtained in this manner from a round robin type experiment will be highly conservative in its ability to reject the null hypothesis of the equivalence of list means.

PROCEDURE

In an effort to reduce the sources of experimental and statistical error described above, a series of experiments was conducted, utilizing a new, recorded test. The experimental test used all the Navy phonetic alphabet words in combination with the digits zero through nine, the total being 260 items. The items were of the same general type as used in the previously described multiple choice test except that there were eight distractors for each item. A typical stimulus was "Peter Four." There were nine possible sets of distractors formed by using the two preceding, adjacent, or following phonetic alphabet words or digits next to the phonetic alphabet word-digit combination which comprise the stimulus. The particular set of distractors was picked by a random process. The 260 items were divided randomly into fourteen lists of eighteen items each. (Eight items were necessarily omitted to obtain a complete set of lists.)

It had been the practice in previous intelligibility tests²⁹ to use one signal-to-noise ratio (the strength of the signal compared to the strength of the noise introduced into the test situation) to increase the level of difficulty of the stimuli. A departure from conventional practice for this experiment was the use

²⁹ See, for example, the reports by John W. Black and other experiments from the Purdue Laboratory listed above.

of three different signal-to-noise ratios.

The eighteen stimuli on a page, which constituted a list, were divided into three parts during the tape recording. The first six stimuli were recorded with an arbitrarily determined signal-to-noise ratio; in the next set of six the voice signal was attenuated two db while the masking noise was held constant; and in the last six stimuli the voice signal was four db below the original voice signal strength or two db below the voice signal strength used for the stimuli just preceding.

This changing of the signal-to-noise ratio has the effect of greatly increasing the range of difficulty of the test. Such a change enables a greater range of tasks to be performed in conjunction with the intelligibility tests and insures that there will be greater flexibility of measurement.

Two other multiple choice intelligibility tests were evaluated along with the experimental test which was being devised. They were the twenty-four word multiple choice intelligibility test developed by the Voice Communication Laboratory at Waco, Texas,³⁰ and referred to hereafter as the Waco Test; and the twenty-seven word multiple choice intelligibility test developed by the United States Naval School of Aviation Medicine³¹ and referred to hereafter as the Pensacola Test. The Waco and the Pensacola tests have been used extensively in intelligibility research. Ten lists were chosen randomly from each of the tests to be used in these experiments. From the Waco Test the lists chosen were (in order of presentation) A₃, A₄, A₈, A₁₂, B₅, B₁₁, B₁₂, A₆, B₄, and A₁₀. From the Pensacola

³⁰ C. H. Haagen, "Intelligibility Measurement: Twenty-Four Word Multiple-Choice Tests," *op. cit.*

³¹ John W. Black, "Multiple-Choice Intelligibility Tests," *op. cit.*

Test the lists chosen were (again in order of presentation) C₁, C₂, C₈, C₉, D₁, D₂, D₄, D₆, D₁₀, and D₁₁.

Method. Two sets of experiments were conducted to aid in constructing and evaluating the experimental test now known as the Purdue Voice Communication Laboratory Intelligibility Test, hereafter referred to as the Purdue Test, and to test the equivalence of the list means within the Waco Test and the Pensacola Test with the more powerful experimental designs.

Experiment One was performed to obtain data concerning the difficulty level of the Purdue and Waco test items, the relationship of the Purdue Test with the Waco Test, the interrelationship of the Purdue and Waco test lists with three environments used in the experimentation, the equivalence of the means of the Purdue Test lists and the equivalence of the means of the Waco Test lists.

Experiment Two was performed in two parts:

1. A study to obtain data concerning the equivalence of the means of ten randomly selected lists from the Pensacola Test along with other information obtained from the design concerning the interrelationships of the lists and listener groups.
2. Investigations of a revised Purdue Voice Communication Laboratory Intelligibility Test.

The Purdue test was revised on the basis of data obtained in Experiment One. The average difficulty had been computed for each item, and the items rearranged among the lists until the means of the lists were theoretically within .35 of one per cent of one another.

These revised tests were recorded and played back to seven groups of listeners to obtain data to evaluate the equivalence of the means of the revised lists and interrelationship of the lists with the talkers and listening groups.

In Experiment One 62 male subjects from the Purdue R.O.T.C. unit were used as listeners, and one talker was used.

In Experiment Two 124 students in introductory speech courses at Purdue were used as listeners and two talkers were used.

Space does not permit the thorough reporting of the rather complex designs and results of both of the two experiments which comprised this investigation. Hence attention in this report will be concentrated on Experiment Two, with a brief summary of Experiment One included.³²

Instrumentation. Briefly, the instrumentation for recording the intelligibility lists for both experiments consisted of an Altec 21-B microphone system, a Presto 900-A7 amplifier unit, and the pre-amplifier of a Stancil-Hoffman tape recorder. Noise from a Grason-Stadler Random Noise Generator was recorded simultaneously with the speech signal.

For playback of the recorded verbal stimuli, a Presto type 900R1 recorder-reproducer, Navy Device 8-I (Portable Inter-phone Trainer), and ANB-H-1A earphones were used in one of the listening environments of Experiment Two. Appropriate amplifier and loudspeaker systems were used for the other two listening environments of Experiment One.

Administration. In Experiment One the 62 subjects were divided into three groups, and each group assigned to one of the three listening-in-noise environments. The three environments were:

1. A large oval field with no nearby obstructions;

³² Detailed information concerning Experiment One is on file at the Purdue Laboratory.

2. An acoustic chamber treated to break up standing wave patterns and reduce reverberation;
3. A large attic room with hard-surfaced walls and ceiling, with earphones provided for the subjects' reception of the test stimuli.

In Experiment Two, only the large room earphone transmission environment was utilized. Either the first or the second talker, determined randomly, presented the Pensacola Test in the first part or the Revised Purdue Test in the second part of Experiment Two, to the several groups in the presence of the white noise. The identity of the separate listening groups in each of the conditions in Experiment Two was maintained to test the effect of the method of sampling of the groups on other factors in the experiment.

No randomization of presentation of the lists to the listening groups was made. Instructions were given both by tape recorder and on the cover page of the test answer booklets. Any questions pertaining to the procedure to be followed were answered before the testing started.

Design of the Experiments. Four major analyses of variance were computed: One each to analyze the Purdue Test, the Waco Test, the Pensacola

Test, and the Revised Purdue Test. These designs were repeated measurements on the same subjects designs involving nesting, or a factorial classification within a factorial classification. Lists, talkers, groups, and conditions were considered fixed variables in these experiments and listeners were considered a random variable.

RESULTS

Experiment One. As explained previously, detailed results of Experiment One will not be presented here. In brief summary, these were the findings:

1. The Purdue Test produced the same results as the Waco Test.
2. Intelligibility test differences for both Waco and Purdue tests in the three listening environments were non-significant.
3. The interactions of lists for both tests and environments were non-significant.
4. The Purdue Test list means differed from one another at the one per cent point of significance.
5. The Waco Test list means also differed from one another at the one per cent point.

Experiment Two. Table 1 presents the results of the first part of Experiment Two. The two talkers used in Experiment Two were not significantly

TABLE I
ANALYSIS OF VARIANCE OF THE PENSACOLA TEST LISTS EXPERIMENT.

| Source | df | Sum of Squares | Mean Square | F | P |
|--|-----|----------------|-------------|-------|----|
| Talkers | 1 | 50.32 | 50.32 | 1.77 | NS |
| Listener Groups within Talkers | 3 | 108.08 | 36.03 | 1.27 | NS |
| Subjects within Listener | | | | | |
| Groups within Talkers | 60 | 1704.39 | 28.47 | | |
| Lists | 9 | 1504.60 | 167.18 | 27.05 | 1% |
| Lists x Talkers | 9 | 188.95 | 20.99 | 3.40 | 1% |
| Lists x Listener Groups within Talkers | 27 | 202.90 | 7.51 | 1.22 | NS |
| Lists x Subjects within Listener | | | | | |
| Groups within Talkers | 540 | 3335.14 | 6.18 | | |
| Total | 649 | 7098.38 | | | |

different from one another in mean intelligibility as measured by the Pensacola Intelligibility Test lists. The five speech class groups which constituted the listeners did not differ significantly from one another. The source of variance for the ten randomly selected lists was significant at the one per cent point. From Black's conclusions, cited earlier, this result might seem surprising. However, as was shown in the review of previous experiments, the result seemed probable.

The significance at the one per cent point of the "Lists x Talkers" interaction assumes a great importance in the measurement of intelligibility. The interaction was obtained by using a non-confounded design instead of the "round robin" method of obtaining the data. As portrayed in the graph in Figure 1, the differences found across the lists were not constant from talker to talker but differed from list to list depending upon which talker read the

lists. It is here that the crux of the problem of intelligibility measurement lies. Heretofore, some of the constructors of intelligibility tests^{33,34} have selected a population of printed stimuli, distributed them randomly into various groups of equal number, and implied that the resulting printed lists were equivalent in difficulty because of the method of construction. Studies by Haagen and Black, cited earlier, have shown by inappropriate methods that after some reallocation of stimuli the means of the lists will be equivalent. In all of the above studies it was the written stimuli which were defined and randomly distributed. If the trait to be measured was an exclusive function of the written stimuli, then any experiment to determine the equivalence of the means of lists which have been pro-

³³ J. O. Egan, "Articulation Testing Methods," *Laryngoscope*, 58 (1948), 955-991.

³⁴ Harvey Fletcher and J. C. Steinberg, "Articulation Testing Methods," *Bell System Tech. J.*, 8 (1929).

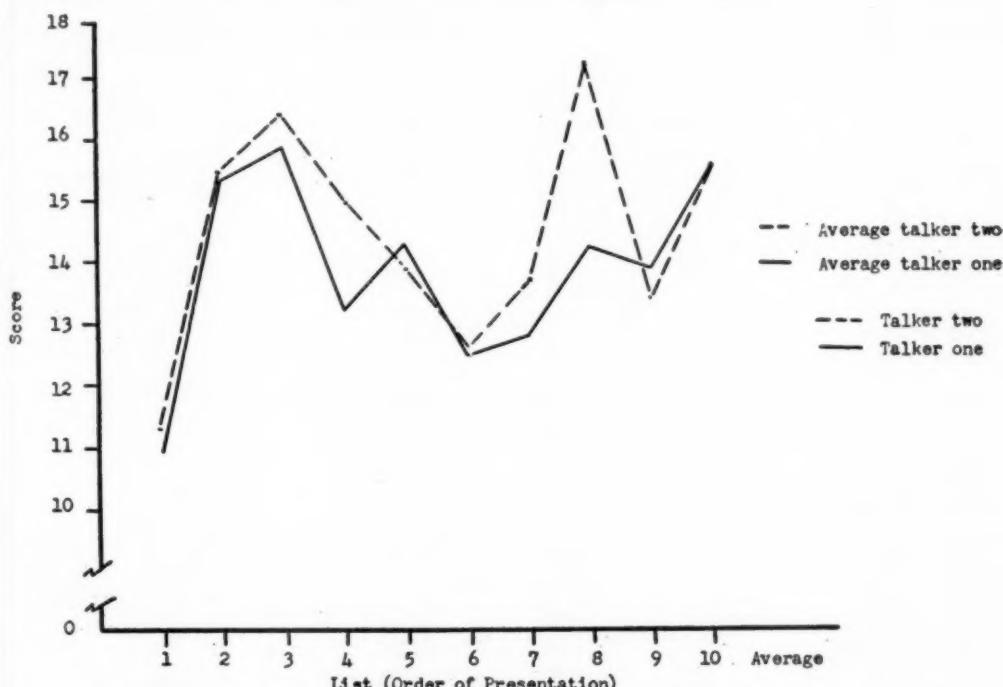


FIG. 1—Graphical Representation of Lists x Talkers' Interaction. Pensacola Test Data.

TABLE II
ANALYSES WITHIN TALKERS FOR THE PENSACOLA TEST.

| Source | Talker One | | | F | P |
|---|------------|----------------|--------------|-------|----|
| | df | Sum of Squares | Mean Squares | | |
| Listener Groups | 2 | 74.74 | 37.37 | 1.52 | NS |
| Subjects within Listener Groups | 32 | 788.96 | 24.66 | | |
| Lists | 9 | 786.96 | 85.44 | 14.17 | 1% |
| Lists x Listener Groups | 18 | 108.77 | 6.04 | 1.00 | NS |
| Lists x Subjects within Listener Groups | 288 | 1735.46 | 6.03 | | |
| Total | 349 | 3494.89 | | | |
| Talker Two | | | | | |
| Listener Groups | 1 | 33.34 | 33.34 | 1.03 | NS |
| Subjects within Listener Groups | 28 | 919.25 | 32.83 | | |
| Lists | 9 | 924.59 | 102.73 | 16.18 | 1% |
| Lists x Listener Groups | 9 | 94.13 | 10.45 | 1.65 | NS |
| Lists x Subjects within Listener Groups | 252 | 1599.68 | 6.35 | | |
| Total | 299 | 3570.99 | | | |

duced by a supposedly random process would merely be a testing of the randomness of the method of construction. Hence, it would seem that the equivalence of the means of some measure of the printed words of lists is not the same as saying that the means of the lists of words spoken from these lists are the same.

This hypothesis gains further credence from the next line in the analysis of variance given in Table I. The "Lists x Listener Groups within Talkers" interaction was not significant. This means that although the means of the recorded lists differed among themselves, the differences remained constant from listener group to listener group as

TABLE III
LISTS X SUBJECTS WITHIN LISTENER GROUPS WITHIN TALKERS:
SUB-ANALYSES OF VARIANCE FOR THE PENSACOLA TEST LIST EXPERIMENT.

| Source | Listener Groups within Talker One | | | F | P |
|------------------|-----------------------------------|----------------|-------------|------|----|
| | df | Sum of Squares | Mean Square | | |
| Lists | 9 | 175.04 | 19.45 | 2.64 | 1% |
| Subjects | 9 | 400.64 | 44.52 | | |
| Lists x Subjects | 81 | 597.56 | 7.38 | | |
| Total | 99 | 1173.24 | | | |
| Group Two | | | | | |
| Lists | 9 | 414.70 | 46.07 | 7.68 | 1% |
| Subjects | 14 | 245.90 | 17.56 | | |
| Lists x Subjects | 126 | 756.50 | 6.00 | | |
| Total | 149 | 1417.10 | | | |
| Group Three | | | | | |
| Lists | 9 | 288.00 | 32.00 | 6.81 | 1% |
| Subjects | 9 | 142.60 | 15.84 | | |
| Lists x Subjects | 81 | 381.40 | 4.70 | | |
| Total | 99 | 812.00 | | | |

the tapes were replayed. Thus, as is shown in Table 2, the tape recorded voice signals comprising the lists were perceived in the same manner from listener group to listener group within talker groups.

Tables 3 and 4 show the individual group breakdowns for the five groups of listeners. These data show that the source of variance for the ten Pensacola lists was significant in all five of the listening groups. Table 3 shows the analyses for the three groups who listened to Talker One speak the lists, while Table 4 shows the analyses for the two groups who listened to Talker Two. The list means for each listener group,

for each talker, and for the total experiment, and the mean intelligibility for each listener group and each talker are presented in Table 5.

Table 6 presents the results of the second part of Experiment Two, that in which the revised Purdue Voice Communication Laboratory Intelligibility Test was evaluated. The two talkers used in this experiment were the same as those used in the first part of Experiment Two, the evaluation of the Pensacola Test. Students in seven speech class groups were used as listeners.

In this experiment the two talkers were significantly different from each other in mean intelligibility. The

TABLE IV
LISTS X SUBJECTS WITHIN LISTENERS GROUPS WITHIN TALKERS:
SUB-ANALYSES OF VARIANCE FOR THE PENSACOLA TEST.

| Listener Groups within Talker Two | | | | | |
|-----------------------------------|-----|----------------|-------------|-------|----|
| Group One | | | | | |
| Source | df | Sum of Squares | Mean Square | F | P |
| Lists | 9 | 376.69 | 41.85 | 6.89 | 1% |
| Subjects | 14 | 407.69 | 29.12 | | |
| Lists x Subjects | 126 | 765.11 | 6.07 | | |
| Total | 149 | 1549.49 | | | |
| Group Two | | | | | |
| Lists | 9 | 642.03 | 71.33 | 10.77 | 1% |
| Subjects | 14 | 511.56 | 3.65 | | |
| Lists x Subjects | 126 | 834.57 | 6.62 | | |
| Total | 149 | 1988.16 | | | |

TABLE V
TABLE OF MEANS FOR THE PENSACOLA TEST LISTS EXPERIMENT.

| Listener Group | Talker One | | | | | | | | | | |
|----------------|------------|------|------|------|------|------|------|------|------|------|-------|
| | Talker Two | | | | | | | | | | |
| 1 | 10.5 | 15.4 | 16.4 | 13.8 | 14.1 | 13.0 | 12.8 | 14.3 | 13.3 | 16.4 | 14.0 |
| 2 | 11.3 | 16.4 | 16.4 | 14.1 | 14.1 | 12.4 | 13.0 | 14.3 | 15.5 | 16.1 | 14.4 |
| 3 | 10.9 | 13.7 | 14.7 | 11.9 | 14.9 | 12.4 | 12.8 | 13.9 | 12.4 | 15.0 | 13.3 |
| | 10.9 | 15.3 | 15.9 | 13.4 | 14.3 | 12.5 | 12.8 | 14.2 | 13.9 | 15.8 | 13.95 |
| 4 | 12.1 | 15.4 | 16.2 | 14.1 | 13.1 | 12.7 | 13.7 | 16.8 | 12.2 | 15.3 | 14.2 |
| 5 | 10.5 | 15.5 | 16.7 | 15.9 | 14.8 | 12.6 | 13.7 | 18.1 | 14.6 | 16.1 | 14.8 |
| | 11.3 | 15.5 | 16.4 | 15.0 | 13.9 | 12.6 | 13.7 | 17.4 | 13.4 | 15.7 | 14.51 |
| Grand Average | 11.1 | 15.4 | 16.2 | 14.1 | 14.2 | 12.6 | 13.3 | 15.7 | 13.7 | 15.8 | 14.26 |

TABLE VI
ANALYSIS OF VARIANCE OF THE REVISED PURDUE TEST LISTS EXPERIMENT.

| Source | df | Sum of Squares | Mean Square | F | P |
|--|-----|----------------|-------------|-------|----|
| Talkers | 1 | 584.25 | 584.25 | 45.64 | 1% |
| Listener Groups within Talkers | 5 | 89.36 | 17.87 | 1.40 | NS |
| Subjects within Listener Groups within Talkers | 52 | 665.48 | 12.80 | | |
| Lists | 12 | 771.17 | 64.26 | 23.23 | 1% |
| Lists x Talkers | 12 | 345.93 | 28.83 | 10.33 | 1% |
| Lists x Listener Groups within Talkers | 60 | 168.72 | 2.81 | 1.01 | NS |
| Lists x Subjects within Listener Groups within Talkers | 624 | 1743.31 | 2.79 | | |
| Total | 766 | 4368.22 | | | |

"Listener Groups within Talkers" mean squares were not unequal to one another.

The means of the lists of the revised Purdue Test differed among themselves, a difference which was significant at the one per cent point. Again, the "Lists x Talkers" interaction was significant. The plot of this interaction is depicted in Figure 2. It demonstrates, as did Figure 1, that the differences from list to list were not constant across "Talkers"

despite all the precautions taken to produce the sounds from the printed stimuli in the same manner from list to list. The "Lists x Listener Groups within Talkers" interaction was not significant.

The list means for each listener group, for each talker, and for the total experiment as well as the listener groups means and talker means are given in Table 7.

It would seem that the proposal of

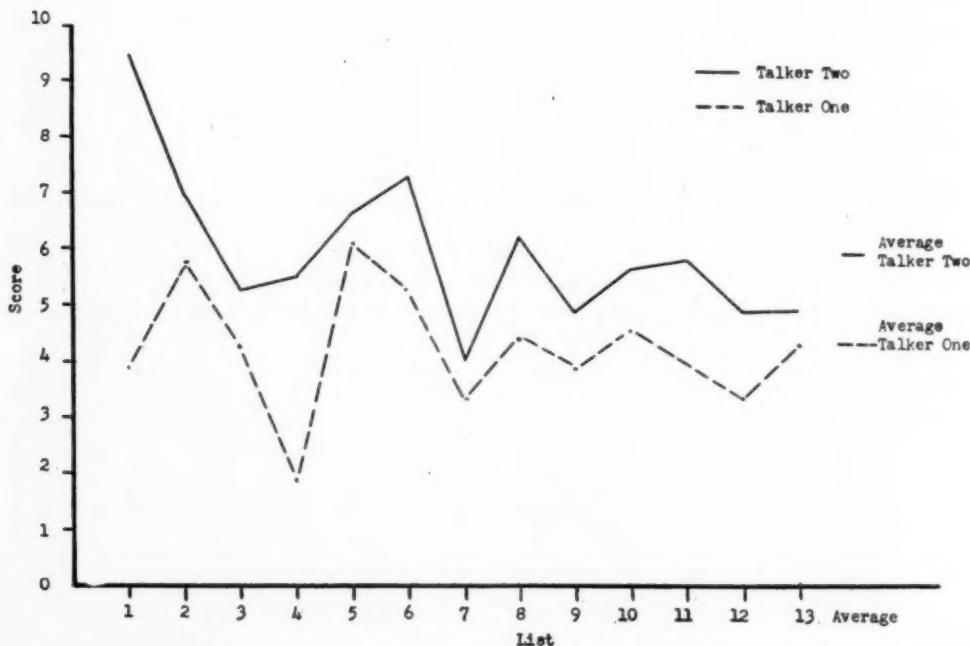


FIG. 2—Graphical Representation of Lists x Talkers' Interaction. Revised Purdue Test Data.

TABLE VII
TABLE OF MEANS FOR THE REVISED PURDUE TEST EXPERIMENT.

| Listener Group | Talker One | | | | | | | | | | | | | Average |
|-----------------------|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|---------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | |
| 1 | 4.8 | 6.0 | 4.8 | 2.0 | 6.2 | 5.7 | 4.0 | 4.2 | 3.9 | 5.3 | 3.9 | 3.9 | 4.7 | 4.6 |
| 2 | 3.8 | 5.8 | 4.4 | 2.1 | 7.1 | 5.8 | 3.3 | 5.0 | 4.2 | 4.4 | 4.6 | 3.4 | 5.2 | 4.6 |
| 3 | 3.1 | 5.5 | 3.6 | 1.4 | 5.2 | 4.5 | 2.7 | 4.1 | 3.6 | 4.0 | 3.5 | 2.7 | 3.1 | 3.6 |
| Talker One Average | 3.9 | 5.8 | 4.3 | 1.8 | 6.1 | 5.3 | 3.4 | 4.4 | 3.9 | 4.6 | 4.0 | 3.4 | 4.3 | 4.24 |
| Talker Two | | | | | | | | | | | | | | |
| 4 | 9.6 | 5.4 | 5.0 | 7.0 | 6.4 | 6.6 | 4.0 | 5.8 | 4.0 | 5.2 | 5.8 | 4.8 | 6.0 | 5.8 |
| 5 | 10.6 | 7.2 | 4.8 | 6.8 | 7.0 | 7.6 | 4.2 | 5.6 | 6.2 | 6.8 | 5.4 | 4.8 | 5.6 | 6.4 |
| 6 | 9.6 | 8.8 | 4.4 | 4.4 | 6.8 | 7.4 | 3.4 | 6.2 | 4.4 | 6.0 | 6.8 | 3.4 | 4.8 | 5.9 |
| 7 | 8.9 | 6.8 | 5.8 | 5.0 | 6.5 | 7.4 | 7.3 | 6.6 | 4.9 | 5.3 | 5.7 | 5.5 | 4.9 | 6.0 |
| Talker Two Average | 9.4 | 7.0 | 5.3 | 5.5 | 6.6 | 7.3 | 4.1 | 6.2 | 4.9 | 5.6 | 5.8 | 4.9 | 5.2 | 5.99 |
| Overall Average | 6.7 | 6.4 | 4.8 | 3.7 | 6.4 | 6.3 | 3.7 | 5.3 | 4.4 | 5.1 | 4.9 | 4.1 | 4.7 | 5.13 |

this experiment to equate the means of a series of listening intelligibility lists was not accomplished. However, it will be noted that from the results of the evaluation of the Pensacola Test lists, a new concept of the equality of intelligibility lists has evolved. The phenomenon of intelligibility is essentially one based on the meaningfulness of the human voice. This is a phenomenon of perception of spoken stimuli, not written stimuli. Therefore, it should be the spoken stimuli which are equated for listening lists, not the written stimuli, and it will be the recordings of the lists which will have been demonstrated to be equivalent, not the written lists. This is especially true since it has been demonstrated that each talker reads each of the lists in a manner which is inconsistent, at least to some degree, with respect to intelligibility. For any talker, it is already known that he usually will be different by a constant amount from any other talker, but that this difference is not constant for all lists has not been specified by previous investigators. Thus, the only way to demonstrate that a series of lists in a listening intelligibility test are not un-

equal in mean intelligibility is to record the lists, play them back to several groups of listeners, and select the recordings of those lists which have been shown by empirical methods to be equivalent. With tape recordings this is not a difficult task to accomplish.

To demonstrate that the means of the lists are not unequal, a method of comparing means within a main effect of an analysis of variance must be used. One of the recent and simple methods of solving this problem has been developed by Scheffe.³⁵ Differences between all possible pairs of means can be tested to discover true differences between them. In these tests the significance level applies to the whole set of tests rather than to any one. Thus, all the differences and equalities are correct, or one or more is incorrect. The significance level in this case is five per cent.

Tables 8 and 9 show the absolute differences between any two means. Significance is indicated when the absolute difference exceeds the confidence

³⁵ H. Scheffe, "An Analysis of Variance of Paired Comparisons," *J. Amer. Statist. Assn.*, 47 (1952), 381.

TABLE VIII
ABSOLUTE DIFFERENCE BETWEEN PAIRS OF MEANS

| Talker Two Revised Purdue Test | | | | | | | | | | | | |
|---|------|--|------|------|------|------|------|------|------|------|------|----|
| 13 means 624 df N = 30 MS for error = 2.79 $\gamma(.05) = 3.38$ | | $S_y = \sqrt{\frac{2.79}{30}} = \sqrt{.0930} = .305$ | | | | | | | | | | |
| Confidence limits $(.05)$ about a difference = $3.38 \times .305 = 1.03$ | | | | | | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| 2.46 | 4.16 | 3.90 | 2.80 | 2.13 | 5.36 | 3.20 | 4.56 | 3.80 | 3.60 | 4.53 | 4.26 | |
| 1.70 | 1.44 | .34 | .33 | 2.90 | .74 | 2.10 | 1.34 | 1.14 | 2.07 | 1.80 | | |
| | .26 | 1.36 | 2.03 | 1.20 | .96 | .40 | .36 | .56 | .37 | .10 | | |
| | | 1.10 | 1.77 | 1.46 | .70 | .66 | .10 | .30 | .63 | .36 | | |
| | | | .67 | 2.56 | .40 | 1.76 | 1.00 | .80 | 1.73 | 1.46 | | |
| | | | | 3.23 | 1.07 | 2.43 | 1.67 | 1.47 | 2.40 | 2.13 | | |
| | | | | | 2.16 | .80 | 1.56 | 1.76 | .83 | 1.10 | | |
| | | | | | | 1.36 | .60 | .40 | 1.33 | 1.08 | | |
| | | | | | | | .76 | .96 | .03 | .30 | | |
| | | | | | | | | .20 | .73 | .46 | | |
| | | | | | | | | | .93 | .66 | | |
| | | | | | | | | | | .27 | | |

limits about the difference. In Table 8 the value which must be exceeded for significance is 1.03. Table 8 is for Talker Two and Table 9 is for Talker One.

From these tables it can be seen, for instance, that for Talker Two lists 3, 4, 8, 9, 10, 11, 12, and 13 could not be

considered unequal at an overall five per cent level of significance. Other groups of lists could have been selected in a similar manner. Another group which could not be considered unequal is 1, 3, 7, 9, 11, 12, and 13 for Talker One.

TABLE IX
ABSOLUTE DIFFERENCES BETWEEN PAIRS OF MEANS

| Talker One Revised Purdue Test | | | | | | | | | | | | |
|---|------|--|------|------|------|------|------|------|------|------|-----|----|
| 13 means 624 df N = 29 MS for error = 2.79 $\gamma(.05) = 3.38$ | | $S_y = \sqrt{\frac{2.79}{29}} = \sqrt{.0962} = .310$ | | | | | | | | | | |
| Confidence limits $(.05)$ about a difference = $3.38 \times .310 = 1.05$ | | | | | | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| 1.86 | .35 | 2.10 | 2.21 | 1.38 | .58 | .48 | .03 | .66 | .04 | .58 | .38 | |
| 1.51 | 3.96 | .35 | .48 | 2.44 | 1.38 | 1.89 | 1.20 | 1.82 | 2.44 | 1.48 | | |
| | 2.45 | 1.86 | 1.03 | .93 | .13 | .38 | .31 | .31 | .93 | .03 | | |
| | | 4.31 | 3.48 | 1.52 | 2.58 | 2.07 | 2.76 | 2.14 | 1.52 | 2.48 | | |
| | | | .83 | 2.79 | 1.73 | 2.24 | 1.55 | 2.17 | 2.79 | 1.83 | | |
| | | | | 1.96 | .90 | 1.41 | .72 | 1.34 | 1.96 | 1.00 | | |
| | | | | | 1.06 | .55 | 1.24 | .62 | .00 | .96 | | |
| | | | | | | .51 | .18 | .44 | 1.06 | .10 | | |
| | | | | | | | .69 | .07 | .55 | .41 | | |
| | | | | | | | .62 | 1.24 | .28 | .62 | | |
| | | | | | | | | .62 | .34 | .96 | | |

DISCUSSION AND CONCLUSIONS

It should be emphasized that these two groups of lists are considered to be equated only for the recordings which these two talkers originally made. It might be that a repeated recording of these lists might not achieve the same results, but that the same process of testing the lists with several groups of listeners would have to be repeated to select a set or several sets of lists considered to be equal.

It should be noted that the above discussion is for the equating of listening lists only. To equate a series of printed or written stimuli, which when read by a speaker or a number of speakers and considered to be equated from the results of subjects listening to them, is a different problem. It is also one which is important in practical work. It is quite likely that the number of situations in which it is desired to test various conditions for speaking, such as in gas masks while performing various other tasks, or to determine the correlates of intelligibility, is as great as the number of situations in which it is desired to evaluate listening conditions or systems using equated tape recorded stimuli.

For the speaking evaluation situation, many randomly chosen speakers should be selected because of the now known speaker-list interaction. Secondly, each of these speakers should speak the stimuli to be used in the lists to several groups of listeners. Next, an appropriate error term will have to be computed. Since the design will require that several talkers read the lists to at least one group, this computation will constitute a problem because of the random variable (talkers) in the repeated measurements over the groups of listeners.

When printed lists which cannot be

shown to be different have been selected for the use of talkers in this manner, an experiment run with them will not be as efficient as a similar listening experiment using tape recorded spoken lists, selected in the manner described previously. This is because of the known variability of a given talker reading a series of lists, the printed stimuli of which have been randomly distributed.

For these reasons only the method and no actual stimulus lists or recordings are included in this report. It is felt that the type of item and the method of construction of the lists and the methods of analysis of the results are more useful than written lists of words.

One other point of interest which may have been demonstrated in these data is an adaptation effect which possibly occurs during the reception of the stimuli of the first list of a test. In the Pensacola Test data, Experiment Two, the first list presented is consistently low for all listening groups regardless of the talker (Table 5). The Revised Purdue Test data, Experiment Two, give inconsistent results. For Talker Two, the first list had the highest intelligibility of any of the lists, while for Talker One the mean intelligibility was below average. However, three other lists had lower means. (Figure 2).

From these results, it would seem that for greater precision in experimental work using intelligibility tests, it would be wise to have the subjects actually practice taking an intelligibility test list besides giving or having them read the instructions.

Within the limitations of this study the results seem to justify the following conclusions:

1. In the experiment involving three environments, it must be inferred that each list mean of the original Purdue Test relative to all other list means re-

mains essentially the same across all three environments.

2. The means of the lists of the Waco Intelligibility Test differ among themselves.

3. In the experiment involving three environments, it must be inferred that each list mean of the Waco Test relative to all other list means remains essentially the same across all three environments.

4. The means of the lists of the Pensacola Intelligibility Test differ among themselves.

5. The means of the Pensacola Intelligibility Test lists change relative to one another as different speakers speak the lists.

6. It must be inferred that randomly distributing stimuli for equated printed lists is not the same as randomly distributing the spoken responses from these stimuli; and therefore, printed intelligibility lists cannot be considered to be equal for experiments which involve listening intelligibility.

7. It must be inferred that the mean of each recorded intelligibility list, relative to all other recorded intelligibility

lists means, remains essentially the same from listening group to listening group.

8. The means of the lists of the revised Purdue Voice Communication Laboratory Intelligibility Test differ among themselves.

9. The means of the Purdue Intelligibility Test lists change relative to one another as different speakers speak the lists.

10. These experiments with the Purdue Intelligibility Test demonstrate that:

- a. Intelligibility lists can be constructed without reference to previous test material.
- b. The nine-celled answer item with random distractors is feasible for intelligibility testing.
- c. The playback of recordings of intelligibility test stimuli is a practical method of administering listening intelligibility tests.
- d. The changing of the signal-to-noise ratio for various sections of items within a list is a practical method of spreading the range of measurement for intelligibility tests.

[au] MONOSYLLABLES

JAMES W. ABEL
Brooklyn College

LIKE an earlier study by the writer¹ the investigation reported here is concerned with phonetic context and diction source material. This investigation differs from the previous one in being more limited in scope and from other studies of phonetic context in that it considers not only the sound which appears on one side or the other of a given sound but considers together the sounds which appear on each side of a given sound. Briefly, it is chiefly concerned with bilateral rather than unilateral relationships.

As a part of a larger study,² Malone has specified the consonants and consonant clusters which precede and follow [au] in monosyllables of the CVC (consonant, or consonant cluster, + syllabic + consonant, or consonant cluster) type. The purpose of the present study is to scrutinize further the phonetic structure³ of this type of [au] monosyllable by showing the pairs of consonant units between which [au] appears in a list of CVC type monosyllables.

As used here, *consonant unit* refers to a consonant or a consonant cluster. For example, both [p] and [pl] are consonant units. Consonant units grouped together because they have the same initial sound constitute a *sound*

type. Consequently, for example, units beginning with [p] constitute a sound type, and the members of this sound type are [p] *type units*. Consonant units grouped together because their initial elements have the same place of articulation constitute a *place type*. Therefore, for example, all units beginning with [t] constitute an *alveolar type*. Each alveolar type is in turn made up of a number of alveolar units, an *alveolar unit* being any unit beginning with an alveolar consonant; similarly, *dental unit* and the like.

When applied to clusters, the classification by first sound, as outlined in the preceding paragraph, may be considered inconsistent, because it identifies following clusters in terms of consonants immediately adjacent to the syllabic while it identifies the preceding clusters in terms of consonants not immediately adjacent to the syllabic. However, the system followed in this analysis has two practical advantages. It identifies clusters in a familiar manner, i.e., by first elements, and its use makes the present findings more directly comparable to those of a prior related study. Therefore, so long as the possible inconsistency of classification by first element is recognized, its practical advantages appear to justify its use.

Subsequently often referred to simply as *the monosyllables*, the raw data for this study is a list of CVC type monosyllables containing the syllabic [au]. The list was compiled by making a word-by-word survey of Kenyon and Knott's *A Pronouncing Dictionary of American English* and Thorndike and Lorge's *The Teacher's Word Book*

¹ James W. Abel, "The Phonetic Contexts of [ɔɪ]," *Speech Monographs*, XX (1953), 247-52.

² Kemp Malone, "The Phonemic Structure of English Monosyllables," *American Speech*, XI (1936), 203-18.

³ There being no concern here with non-distinctive variants, to the degree that non-distinctive is synonymous with phonemic, this might be looked upon as a study of phonemic structure. The underlying assumptions are, however, those of traditional phonetics.

TABLE I

| Occurring Initial Units | | | | | Occurring Final Units | | | | |
|-------------------------|----|-----|----|----|-----------------------|------|-----|-----|-----|
| p | v | s | z | k | p | ð | n | l | f |
| pl | θ | sp | zw | kl | pt | ðd | nz | ld | ſt |
| pr | ð | spr | r | kr | ps | ðz | nt | lz | tſ |
| b | t | st | ſ | g | f | t | nts | lθ | tſt |
| bl | tr | str | ſr | gl | ft | ts | nd | lθs | dʒ |
| br | d | sk | tſ | gr | fs | d | ndz | s | dʒd |
| m | dr | skr | dʒ | h | θ | dz | ns | st | k |
| w | n | sm | j | hw | θt | nst | sts | kt | |
| f | l | sn | | | θs | ndʒ | z | ks | |
| fl | | sl | | | | ndʒd | zd | ŋ | |
| fr | | sw | | | | | | | |

of 30,000 Words and by carefully scanning *Webster's New International Dictionary, Second Edition*.⁴ Except for proper nouns and adjectives and words labelled "foreign" in *Webster's*, all root forms of the CV and CVC types containing [au], which were found, were listed. Subsequently, the list was expanded and completed by adding the inflected forms of the recorded roots. This converted the CV type roots into CVC type monosyllables and added to the list CVC type monosyllables with more complex termini.

The focal purpose being to show the pairs of consonant units between which [au] appears in this list of monosyllables, eventually the list must be analyzed in terms of the pairings of consonant units found. However, a helpful preliminary step is to set down without reference to their pairings two unilateral structural sets,⁵ one consisting of the consonant units which precede [au] and the other consisting of the

⁴ For the reasons for the use of such a source list and for details of its nature see "Phonetic Contexts of [ɔɪ]," p. 247.

⁵ "A STRUCTURAL SET is a group of all the phonemes which occur in a phonetic environment. . . ." Bernard Bloch and George L. Trager, *Outline of English Analysis* (Baltimore, 1942), p. 45.

consonant units which follow [au] in CVC type monosyllables. Respectively, these are referred to in this report as *occurring initial units* and *occurring final units* and are listed in Table I.

Since the present study is somewhat of an extension of work done already by Malone, and since the findings given in Table I are of the same nature as his,⁶ it seems appropriate to compare the two sets of results. They agree except as stated below. Before this statement is made, the manner in which the writer has dealt with Malone's findings should be described. Since his findings about the consonant units which precede and follow [au] are stated in essentially negative terms, to transpose his results into positive terms, as it appears desirable to do, it has been necessary to check his negations about [au] monosyllables against his positive findings about initial and final consonant units in monosyllables of all kinds. It is on the basis of such checks that the following positive findings are attributed to him.

In effect, he shows all the occurring initial consonant units which appear in

⁶ "The Phonemic Structure of English Monosyllables," pp. 211-18.

Table I except [zw], [θ], [skr], and [hw]. The absence of [zw] may be due to the fact that Malone's analysis of existent initial English clusters does not include [zw].⁷ In fact, a list of initial English clusters composed by the writer from various cluster analyses does not contain [zw]. For the circumstances under which initial [zw], [θ], [skr], and [hw] occur here, see the subsequent sections dealing with [z], [θ], [s], and [h] type units.

Turning to occurring final units, on the one hand, Malone found all those recorded here except the [p], [f], [ʃ], [k], and [ŋ] units given in Table I. For the details of the occurrence of final [ŋ], see the subsequent section dealing with [t] type units; otherwise, see the sections devoted to [p], [f], [ʃ], and [k] type units. On the other hand, Malone appears to say that certain units which have not been found in the present investigation follow [au]. They are [dθ], [tst], [dθs], [ndz], and [ntst]. It is not possible to verify this deduction by reference to examples, because none are given.

The following sections give an expanded answer to the question raised by the specific purpose: What are the pairs of consonant units between which [au] appears in the monosyllables? Since of necessity there is much detail in these sections, and since ease of reference is desired, it appears advisable to describe the arrangement of this material. The first sections deal with bilabial types, followed in turn by sections devoted to labio-dental, dental, alveolar, palato-alveolar, palatal, velar, and glottal types. Though in a few cases the nature of the material dictates deviation, characteristically, each section consists of two paragraphs deal-

ing with a single sound type, [p] type units for example. The first paragraph, which views the pairings from the standpoint of occurring initial clusters, consists of two parts. First, there are summary statements about the nature of the pairings between the initial unit, or units, under consideration and the occurring final units. Second, the specific pairings found in the monosyllables are given and exemplified. These are grouped by initial units, appearing in the same order as do the occurring initial units in Table I. The second paragraph of each section views the pairs found from the standpoint of the occurring final units and states the initial units with which the final units in point pair. To avoid needless repetition, usually, examples are not given in the second paragraphs. With the information provided by the second paragraphs, it is always possible to find desired examples by reference to the pertinent first paragraphs.

Except for the pairing of [p] with [tʃ(t)], the occurring initial [p] units pair only with final alveolar units, and further, one or more initial [p] unit pairs with at least one unit of each of the six final alveolar types. In other words, at least one of the initial [p] type units pairs with at least one occurring final [t], [d], [n], [l], [s], and [z] type unit. Within this larger pattern, all occurring initial [p] units appear opposite final [z].

[p-] pairs with [t(s)], [nd(z)], [ns(t)], [s], [z], [tʃ(t)]: *pout(s)*,⁸ *pound*, *pounce*, *pouce*, *pows*,⁹ colloq., *pouch*.

⁸ Subsequently, when transcriptions clearly indicate the inclusion of inflected forms, only pertinent root forms will appear among the examples.

⁹ + following an example means there is at least one other word containing the same pair of units. Otherwise, unless so stated, the words listed are the sole ones containing the context under consideration.

⁷ Kemp Malone, "The Phonemes of Current English," in *Studies for William A. Read* (Louisiana State University Press, 1940), p. 153.

[pl-] pairs with [t(s)], [d], [z], [ns(t)]: *plout*, Scot., *plow-ed*,^{s,10} *plounce*, obs.¹¹

[pr-] pairs with [d], [dz], [l(d)(z)], [z]: *proxed*+, *prouds*, arch., *prowl*, *prows*+

The three final [p] type units which occur after [au] are paired in dialect words with initial [d], [k], [l], [sk], and [r], except final [pt] with initial [d].

Except for the pairing of initial [b] with final [k(t)(s)], the occurring initial [b] type units parallel the initial [p] type units almost exactly. They have the same structural pattern, and with the exception already given, they also pair only with final alveolar units. Furthermore, one or more of the initial [b] type units pairs with at least one unit of each of the alveolar types. Also, all three of the occurring initial [b] units pair with final [z(d)].

[b-] pairs with [t(s)], [d], [dz], [n(z)], [nd(z)], [ns(t)], [l(d)(z)], [s(t)], [z(d)], [k(t)(s)]: *bout*+, *boughed*+, *bouds* [bau(u)dz], obs., *boun* [bau(u)n], arch., *bound*+, *bounce*, *bowel* [bau(ə)l], *bouse*, naut., *bowze* [bu(au)z]+, *bowk* [bau(o)k]+, bleaching.

[bl-] pairs with [t(s)], [s], [z(d)]: *blout*, Scot., *blouse*, n., *blouse*+, v.

[br-] pairs with [d], [n(d)(z)], [st(s)], [z(d)]: *browed*, *brown*, *browst* [brau(u)st], dial., *browse*+

Neither final [b] nor [m] type units follow [au] in the monosyllables.

Initial [m] alone does occur before [au], however, as does the remaining bilabial, [w]. [m] before [au] appears opposite all the final dental units and

¹⁰ Orthographic forms like this one indicate exemplification of two phonetic contexts by inflected forms of the same root.

¹¹ Many of the words labelled "obsolete" survive as dialectal forms.

certain of the final alveolar units, but not opposite any [l] type unit and opposite only one [t] type unit, this in an obsolete inflected form of *may*.

[m-] pairs with [θ(t)(s)], [ð(z)(d)], [t], [d], [z], [nt(s)], [nd(z)], [s], [zd]: *mouth*, n. and v., *mouth* [mau(o)t], obs., *mow-ed*,^{s+}, *mount*, *mound*, *mouse*, n., *moused*.

[w-] pairs with [d], [z], [nd], [ndz]: *wow-ed*,^s, slang and technical, *wound*, *wounds*, interj.

The occurring initial [f] type units have the same structure as the occurring initial [p] and [b] type units, and like them, also form pairs with at least one unit of each of the final alveolar types. In fact, initial [f] units appear opposite final alveolar units only.

[f-] pairs with [d(z)], [nt(s)], [nd(z)], [l(d)(z)]: *foud*, *fount*, *found*+, *foul*+

[fl-] pairs with [t(s)], [ns(t)]: *flout*, *flounce*.

[fr-] pairs with [n(d)(z)], [ns], [nst], [l(z)], [st(s)], [z], [zd]: *frown*, *frounce*+, n., *frounced*, obs., *frowl* [frau(o)l], local, *frowst*+, *frows*+, *frowzed*, obs.

The occurring final [f] type units are like the occurring final [p] type units in that they are found in dialect words. Each of the [f] units is mated with initial [d], [j], [g], and [h].

Initial [v] before [au] is followed by certain alveolar units and [tʃ(t)].

[v-] pairs with [d], [z], [l(d)(z)], [tʃ(t)]: *vow-ed*,^s, *vowel* [vau(ə)l], *vouch*.

Final [v] type units, as final [b] and [m], do not come after [au] in the monosyllables.

Initial [θ] and [ð] occur in [au] monosyllables of limited use and before a limited number of following alveolar units.

[θ-] pairs with [z] in the slang word *thous*.

[ð] pairs with [d], [z], [l(z)]: *thou-ed-s*, *dhoul*, Anglo-Irish.

The occurring final [θ] type units pair with initial [m], [dr], [s], and [r]; the occurring final [ð] type units, with [m] and [s].

The combination of a diphthong + [ŋ] is generally considered "forbidden" in English,¹² but in the loan word *taun* [taʊŋ], final [ŋ] follows [au] and is paired with [t]. Otherwise, the occurring initial [t] type units pair with final alveolar units, but with no [d] type unit and only one [s] type unit, this in a dialectal word. Unlike any of the initial units previously discussed, [t] pairs with itself. *

[t-] pairs with [t(s)], [n(d)(z)], [l(d)(z)], [z], [zd], [ŋ]: *tout*+, colloq., *town*, *towel* [taʊ(ə)l], *taus* [tɔ(u)z]+, *toused*, dial., *taun* [taʊŋ].

[tr-] pairs with [t(s)], [ns(t)], [l(d)(z)], [s], [z]: *trout*, v., *trounce*+, *trowel* [trəʊ(ə)l], *trouse* [trɔ:s], dial., *trouse* [trəʊ(u)z], obs.

Both the final [t] type units which follow [au] are preceded by initial [p, pl], [b, bl], [t, tr], [d, dr], [k, kr], [g, gl, gr], [n], [l], [f], [sp, spr, st, sk, skr, sn], [ʃ], [tʃ], [r], and [j]. Final [t], but not [ts], pairs with initial [m].

Between them, the occurring initial [d] type units pair with a variety of final place types, namely, all except palato-alveolar types. However, the pairings of initial [d] type units with final bilabial, labio-dental, and velar types are restricted to dialectal or obsolete words. The pairings with final alveolar types are not so limited, and at

least one unit of each of the final alveolar types is paired with one or more of the occurring initial [d] type units. That is, either initial [d] or [dr], or both, pairs with each of the existing alveolars either alone or when initial in a final cluster.

[d-] pairs with [p(s)], [f(t)(s)], [t(s)], [d(z)], [n(d)(z)], [l(z)], [ld], [s], [st], [z], [zd]: *doup* [dau(u)p], Scot., *dow(u)f* [dau(u)f] dial., *doubt*+, *dowd*+, *down*, *dowl(le)*+, *doweled*,¹³ *dou(w)se*+, dial. n., *dou(w)sed*, *d(h)ows*+, *dowsed*.

[dr-] pairs with [θ(s)], [t(s)], [d], [dz], [n(d)(z)], [z], [zd], [k(t)(s)]: *drouth*, *drought*, *drowed* [drəʊ(ə)d]+, dial., *drouds*, Scot., *drown*, *drows*+, dial., *drowded*, *drowk*, obs.

Two final [d] type units, [d, dz], occur after [au]. When [d] follows, the initial units are [pl, pr], [b, br], [m], [w], [f], [v], [ð], [d, dr], [n], [l], [s, str, sk, sl, sw], [r], [ʃ, sr], [dʒ], [j], [k, kl, kr]. Of these, [pr], [b], [f], [d, dr], [str], [r], [ʃ, sr], and [kr, kl] precede when [dz] is final.

Initial [n] and [l] before [au] in the monosyllables are alike in their patterning with final units in that both pair with some unit of each alveolar type, except [l] type units. Hence, [n] pairs with itself. Initial [n] and [l] differ in their over-all patterning with final units in that [n] appears opposite only alveolar units while [l] groups with other types as well, specifically, [p], [tʃ], and [k] type units in dialect words.

¹² See Malone, "The Phonemic Structure of English Monosyllables," p. 213, and Leonard Bloomfield, *Language* (New York, 1933), p. 135.

¹³ Though both *A Pronouncing Dictionary of American English* and *Webster's New International Dictionary* show only ['dəʊəl] for *dowel*, it is the writer's observation that those who have occasion to use the word often say [dəʊəl], just as the commoner word *towel*, for example, is often given a one, rather than a two-syllable pronunciation, as is indicated by the inclusion of [taʊl] in the pronouncing dictionary.

TABLE II
FINAL [n] TYPE UNITS

| Initial Units | n | nt | nts | nd | ndz | ns | nst | nz | ndʒ | ndʒd |
|------------------|---|----|-----|----|-----|----|-----|----|-----|------|
| p | | | | x | x | x | x | | | |
| pl | | | | | | o | o | | | |
| b | o | | | x | x | x | x | x | | |
| br | x | | | x | | | | x | | |
| m | | x | x | x | x | | | | | |
| w | | | | x | x | | | | | |
| f | | x | x | x | x | | | | | |
| fl | | | | | | x | x | | | |
| fr | x | | | x | | x | o | x | | |
| t | x | | | x | | | | x | | |
| tr | | | | | | x | x | | | |
| d | x | | | x | | | | x | | |
| dr | x | | | x | | | | x | | |
| n | x | | | | | | | x | | |
| l | o | | | o | | | | o | x | x |
| s | | | | x | x | | | | | |
| st | | | | o | o | | | | | |
| skr | | | | | | | | x | x | |
| sw | | | | o | o | | | | | |
| z | | | | | o | | | o | | |
| zw | | | | | o | | | o | | |
| r | o | | | x | x | x | o | o | | |
| dʒ | | | | | | x | x | | | |
| k | | x | x | | | | | | | |
| kl | x | | | x | | | | x | | |
| kr | x | | | x | | | | x | | |
| g | x | | | x | | | | x | | |
| gr | | | | x | x | | | | | |
| h | | | | x | x | o | | | | |

[n-] pairs with [t], [ts], [d], [n(z)], [s], [z]: *nowt+*, Scot., *knouts*, *nowed* [nu(au)d], her., noun, *nous* [nu(au)s], *nows+*.

[l-] pairs with [p(t)s], [t(s)], [d], [n(d)(z)], [ndʒ(d)], [s], [z], [st(zd)], [tʃ(t)], [k(t)s]: *loup* [lau(o)(u)p], Scot., *lout+*, *loud+*, *lown*, *dial.*, *lounge*, *louse*, n., *lows+*, *dial.*, *louse*, v., *louch* [lau(u)tʃ], *dial.*, *louk* [lau(u)k], *dial.*.

There are many final [n] type units in English, ten of which follow [au] in the monosyllables, grouped with initial units as shown by Table II. In this table and the others which follow "x" or "o" at the intersection of a row and a column means the presence of a pair, but "o" indicates that the pairing occurs only in a form labelled local, dialectal, obsolete, archaic, or rare.

Of the five occurring final [l] type units, two, [lθ(s)], appear only in the Anglo-Irish word *south*. Each of the remaining three, [l(d)(z)], pair with initial [pr], [b], [f], [v], [t, tr], [d], [s, sk], [r], [dʒ], [j], [k], [g, gr], and [h]. Of these [l(d)] also pair with initial [fr], [ð], and [sn] in localisms.

Since [s] is the first element of more initial English clusters than any other consonant, it is hardly surprising to find more [s] type units than any other type before [au] in the monosyllables. Of the eleven different [s] type units which occur, only two, however, [s] and [sk], pair with a relatively large number of final units. Looked at from the standpoint of the final units appearing, it is seen that [t], [ts], [d], and [z] are the only ones which pair with as many as five of the occurring initial [s] units. In general, the patterning of initial [s] type units with final consonant units is characterized by scatter, a fact which becomes apparent when

the pairings are presented in tabular form as in Table III.

This table shows all of the occurring initial [s] units and all of the final place types with which some initial [s] unit pairs. It does not show any labiodental or velar type final units, because no unit of these types pairs with any initial [s] unit. Further, final bilabial type units appear only with [sk] in a dialect word. In fact, if the pairings limited to obsolete, dialectal, and local words are excluded from consideration, the total number of pairs beginning with an [s] type unit are reduced by approximately one-fourth, and hence the scatter of the occurring pairs is increased. This characteristic of scatter should be kept in mind when the summary observations are made that [s] type units pair with more different types of final units than any other set of initial units, and that one or more of the [s] units pairs with at least one unit of each of the alveolar types. As a result, [s] is another consonant which pairs with itself. Other notable details shown by Table III include: (1) The sole appearance of [ʃ(t)] in the monosyllables, this in an Australian term, (2) The first appearance in the analysis of [dʒ(d)], and (3) The pairing of [s] with all the final dental units. This last is a peculiarity which [s] shares with initial [m].

[s-] pairs with [θ(t)s], [ð(d)(z)], [d], [nd(z)], [l(z)], [ld], [lθ(s)], [s(t)], [z]: *south* [sauθ(ð)], v., *soughed*, *sound+*, *sowl(e)* [sau(u)l]+, *dial.*, *sowled* [sau(o)ld], *dial.*, *sowlth*, Anglo-Irish, *souse+*, v., *sows+*.

[sp-] pairs with [t(s)], [s], [z(d)]: *spout*, *spouse* [spaʊz(z)], n., *spoused*, obs.

[spr-] pairs with [t(s)]: *sprout+*.

[st-] pairs with [t(s)], [nd(z)], [ʃ(t)]: *stout+*, *stound* [stau(u)nd], *dial.*, *stoush*, Australian slang.

TABLE III
INITIAL [s] TYPE UNITS

| s | sp | spr | st | str | sk | skr | sm | sn | sl | sw | Final Units |
|---|----|-----|----|-----|----|-----|----|----|----|----|-------------|
| | | | | | o | | | | | | p |
| | | | | | o | | | | | | pt |
| | | | | | o | | | | | | ps |
| x | | | | | | | | | | | θ |
| x | | | | | | | | | | | θt |
| x | | | | | | | | | | | θs |
| x | | | | | | | | | | | ð |
| x | | | | | | | | | | | ðd |
| x | | | | | | | | | | | ðz |
| | x | x | x | | x | o | | x | | | t |
| | x | x | x | | x | o | | x | | | ts |
| x | | | | x | x | | | | x | o | d |
| | | | | x | | | | | | | dz |
| | | | | | | | | | | | n |
| | | | | | | | | | | | nt |
| | | | | | | | | | | | nts |
| x | | | | | o | | | | | o | nd |
| x | | | | | o | | | | | o | ndz |
| | | | | | | | | | | | ns |
| | | | | | | | | | | | nst |
| | | | | | | | | | | | nz |
| | | | | | | x | | | | | ndʒ |
| | | | | | | x | | | | | ndʒd |
| o | | | | | x | | | | | o | l |
| o | | | | | x | | | | | | ld |
| o | | | | | x | | | | | o | lz |
| o | | | | | | | | | | | lθ |
| o | | | | | | | | | | | lθs |
| x | x | | | | x | | x | | | | s |
| x | | | | | | | | | | | st |
| | | | | | | | | | | | sts |
| x | x | | | | x | | x | | x | o | z |
| | x | | | | | x | | x | | | zd |
| | | | | | o | | | | | | ʃ |
| | | | | | o | | | | | | ʃt |
| | | | | | o | x | o | x | | | tʃ |
| | | | | | o | x | o | x | | | tʃt |
| | | | | | | x | | | | | dʒ |
| | | | | | | x | | | | | dʒd |

[str-] pairs with [d(z)]: *stroud*.

[sk-] pairs with [p(t)(s)], [t(s)], [d], [l], [ld], [lz], [s], [z], [tʃ(t)]: *scoup*, dial., *scout*+, *scowed*, *scowl*+, obs. *scowl-ed*, *scouse*, naut., *scows*, *scouch*, dial.

[skr-] pairs with [t(s)], [ndʒ(d)], [dʒ(d)]: *scrount*, dial., *scrounge* [skrau(u)ndʒ], slang, *scrouge* [skru(av)dʒ], colloq.

[sm-] pairs with [s], [z(d)], [tʃ(t)]: *smous(e)*, slang n., *smouse*, v., *smouch*+

[sn-] pairs with [t], [ts], [l(z)], [tʃ(t)]: *snout*+, slang, *snouts*, *snowl*, local, *snouch*, rare.

[sl-] pairs with [d], [z], [tʃ(t)]: *slough-ed*, *s+*, *slouch*.

[sw-] pairs with [d], [nd(z)], [z]: *swowed*, local, *swound* [swau(u)nd], arch., *swows*, dial. var. of *soughs*.

Final [sts] pairs with no initial unit not also paired with final [st], but the reverse is not true. Sometimes [s] and [st] pair with the same preceding units, sometimes not. When [s] is the following unit, the preceding units are [p], [b, bl], [m], [tr], [d], [n], [l], [s, sp, sk, sm], [r], [tʃ], [k], [g, gr], and [h]. Of these, [st] follows [b], [d], [l], [s], [r], [tʃ], and [gr] plus [br], [fr], and [dʒ]. Final [sts] follows when [br], [fr], [r], and [dʒ] precede.

The archaic forms *zounds* and *'swounds* are, respectively, abbreviation and contraction of *God's wounds*. Their varying pronunciations result in the appearance of [z] and [zw] before [au] paired with [n(d)z] as follows:

[z-] with [n(d)z]: *zounds* [zaʊn(d)z].

[zw-] with [n(d)z]:

'*swounds* [z(w)əʊn(d)z].

Final [z] after [au] pairs with more different initial consonant units than any other final unit. Specifically, it appears after all the occurring initial

units except [f, fl], [spr, st, str, skr, sn], [z, zw], [ʃr], [kr], and [gl]. When [zd] follows, the preceding units are [b, bl, br], [m], [fr], [t], [d, dr], [l], [sp, sm], [r], [gr], and [h].

In the monosyllables more different final units combine with initial [r] than any other unit preceding [au]. The pairings of [r] with other than alveolar units occur in dialect words, in one of which it combines with [θ], thus making [r] one of the four initial units pairing with at least one of the occurring final [θ] type units. The other three are initial [m], [dr], and [s]. Within the alveolar group it is found that at least one unit of each final alveolar type pairs with initial [r].

[r-] pairs with [p(t)(s)], [θ], [t(s)], [d], [dz], [n(z)], [nd], [ndz], [ns], [nst], [l(d)(z)], [s], [st(s)], [z], [zd]: *roup* [raʊ(u)p], dial., *routh* [ru(av)θ], dial., *rout*+, *rowed*+, colloq., *rounds*, dial., *roun*, obs., *rounded* [raʊ(u)nd]+, obs., *rounds*, *rounce*+, print., *rounced*, dial., *rowel* [raʊ(ə)l],¹⁴ *rouse*, obs., *roust*+, colloq., *rows*+, dial., *roused*.

The occurring initial [ʃ] type units are limited in their pairings to final alveolar units.

[ʃ-] pairs with [t(s)], [d(z)], [z]: *shout*+, *showd*, Scot., *shou(w)s*.

[ʃr-] pairs with [d(z)]: *shroud*.

The occurring final [ʃ] units, [ʃ(t)], appear only with initial [st] in an Australian slang term.

Initial [tʃ] before [au] is grouped with a few final alveolar units and two velar ones. Only one pairing appears in other than a labelled form.

[tʃ-] pairs with [t(s)], [s], [st], [z], [k(s)]: *chauth*, India, *chiaus*+, obs., *choused*, colloq., *chows*, *chowk*, India.

¹⁴ See footnote 13.

Initial [dʒ] before [au] is followed by alveolar units of all but the [t] type. The pairing with [d] appears in a dialectal form.

[dʒ-] pairs with [d], [ns(t)], [l(z)], [ld], [st(s)], [z]: *jowed* [dʒau(o)d], dial., *jounce, joule* [dʒu(uv)l], *fowled*, obs., *joust* [dʒau(ʌ)(u)st], *jows+*.

Final [dʒ(d)] pairs with initial [skr] and [g].

Initial [j] pairs with the occurring final [f] units in forms of a Scottish word; otherwise, [j] groups with final alveolar units.

[j-] pairs with [f(t)s], [t(s)], [d], [l(d)(z)], [z]: *youff*, Scot., *yowt*, dial., *yowed, yowl, yows+*, n., Scot.

Initial [k] pairs with all the occurring final [tʃ] type units. It does likewise with all occurring final [p] type units, thus becoming the fifth and last unit to do so, the others being [d], [l], [sk], and [r]. All of the pairs ending in a [p] type unit occur in obsolete or dialect words. The remaining pairs begun by a [k] type unit end in an alveolar unit, each of the six final alveolar types being represented. Hence, the [k] type units join with the [p], [b], [f], [d], [s], and [r] type units to form the group of initial types which bears this relation to the final alveolar types.

[k-] pairs with [p(t)s], [d], [nt(s)], [l(d)(z)], [s], [z], [tʃ(t)]: *coup* [kau(o)p], obs., *cowed+*, *count*, *cowl+*, *cous* [kau(u)s], var. of *cowish*, *cows+*, dial., *couch*.

[kl-] pairs with [t], [ts], [d(z)], [n(d)(z)], [z]: *clout+*, obs. n., *clouts, cloud, clown, clows+*.

[kr-] pairs with [t], [ts], [d(z)], [n(d)(z)], [tʃ(t)]: *crout+*, *krauts, slang, crowd+*, *crown, crouch+*.

Final [k] type units are found only in dialectal and local forms and all are paired with initial [dr], [l], [tʃ], [g],

and [h], except that final [kt] is not paired with initial [tʃ].

As is true of the initial [t] type units, the occurring initial [g] type units make some combination with all final alveolar types other than [d] type units. Otherwise and more specifically, all three occurring initial [g] units pair with [t(s)], and one or more of them pairs with at least one unit of each of the remaining alveolar types. In addition, [g, gr] pair with palato-alveolar units. As a result, both initial [g] and [gr] join with initial [p], [v], [l], [sk, sm, sn, sl], and [k, kr] to complete the group of units which pairs with final [tʃ(t)], and initial [g] becomes one of the two units to combine with [dʒ(d)], the other being [skr]. Initial [g] also pairs with the occurring [f] and [k] type units, but only in dialect words.

[g-] pairs with [f(t)s], [t(s)], [n(d)(z)], [l(d)(z)], [s], [z], [dʒ(d)], [k(t)s]: *gowf(f)* [gau(o)f], Scot., *gout+*, *gown, gowl* [gau(u)(o)l]+, dial., *gauss* [gau(ɔ)s], *gaus*, Ger. hist., *gouge, gow(u)k* [gau(o)k], dial.

[gl-] pairs with [t(s)]: *glout* [glu(u)t], arch.

[gr-] pairs with [t(s)], [nd(z)], [l(d)(z)], [s(t)], [z(d)], [tʃ(t)]: *grout+*, *ground, growl, grouse+*, *grouze*, dial., *grouch*.

There is no final [g] type unit, just as there are no final [b], [m], or [v] type units after [au] in the monosyllables.

Except in the New Zealand tree name *whau*, the only occurring initial [h] type unit is [h] alone. It joins with [d], [j], and [g] in pairing with final [f] type units and with [dr], [l], [tʃ], and [g] in combining with final [k] type units, always in local, obsolete, or dialectal forms. Otherwise, [h] pairs with alveolar units.

[h-] pairs with [f(t)(s)], [nd(z)], [ns], [l(d)(z)], [s], [z], [zd], [k(t)(s)]: *howf(f)* [(h)au(o)f], *dial.*, *hound*, *hounce* [(h)auuns], *dial.*, *howl+*, *house*, n., *hows+*, *housed*, *howk* [(h)au(o)k], *obs.*

[hw-] pairs with [z]: *whaus.*

A simplified summary of the preceding sections is given by Table IV. Basically, the simplification results from the fact that each row and each

column of the table represents a type of unit, not an individual unit, except in those cases in which sound type coincides with individual unit. For example, "x" at the intersection of the [f] row and the [d] column means that at least one of the three occurring initial [f] type units pairs with at least one of the two occurring final [d] type units, whereas, "x" at the intersection of the [v] row and the [d] column

TABLE IV
FINAL SOUND TYPES

| Initial Sound Types | p | b | m | f | v | θ | ð | t | d | n | l | s | z | ʃ | ʒ | tʃ | dʒ | k | g | ŋ | |
|---------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|----|----|---|---|---|--|
| p | | | | | | | | x | x | x | x | x | x | x | x | | x | | | | |
| b | | | | | | | | x | x | x | x | x | x | x | | | x | | | | |
| m | | | | | | x | x | o | x | x | | x | x | | | | | | | | |
| w | | | | | | | | x | x | | | x | | | | | | | | | |
| f | | | | | | | x | x | x | x | x | x | x | | | | | | | | |
| v | | | | | | | | x | | x | | x | | x | | | x | | | | |
| θ | | | | | | | | x | | | | x | | | | | x | | | | |
| ð | | | | | | | | x | | | | o | | x | | | | | | | |
| t | | | | | | | | x | | x | x | o | x | | | | x | | | | |
| d | o | o | o | z | | x | x | x | x | x | x | x | x | | | o | | | | | |
| n | | | | | | | x | x | x | | | x | x | | | | | | | | |
| l | o | | | | | x | x | x | | x | x | | | | o | o | | | | | |
| s | o | | | | x | x | x | x | x | x | x | x | x | o | | x | x | | | | |
| z | | | | | | | | | | | | o | | | | | | | | | |
| r | o | | | | o | | x | x | x | x | x | x | x | x | | | | | | | |
| ʃ | | | | | | | x | x | | | | | x | | | | | | | | |
| tʃ | | | | | | | o | | | x | x | | | | | o | | | | | |
| dʒ | | | | | | | | o | x | x | x | x | | | | | | | | | |
| j | | o | | | | | o | x | | x | | x | | | | | | | | | |
| k | o | | | | | | x | x | x | x | x | x | | | | x | | | | | |
| g | | o | | | | | x | | x | x | x | x | | | | x | x | o | | | |
| h | | o | | | | | | x | x | x | x | | | | | | o | | | | |

means that at least one of the two final [d] units pairs specifically with initial [v], there being no other initial [v] type unit.

Expressed in terms of sound types, the occurrences shown by Table IV constitute a bilateral structural set of consonant units which bracket [au] in the CVC type monosyllables studied.

Further summary observations which have been made about the monosyllables follow:

1. There are 47 occurring initial and 46 occurring final units.

2. With a few exceptions, the occurring units are the same as the units found by Malone before and after [au] in CVC monosyllables.

3. There are 398 different pairs made up of preceding and following consonant units.

4. Since there are 47 occurring initial and 46 occurring final units, mathematically there are 2162 possible pairings among the occurring units. Granting this frame of comparison, the actual occurrence of 398 pairs means that only a small proportion of the mathematically possible pairs actually exist.¹⁵

5. All of the English consonants which may begin a word appear initially. This means that there is at least one pair beginning with each of the word-initial consonants and that 22 initial sound types appear.

6. There are 20 consonants which may end a word in English. ([w], [r], [h], and [j] are not considered to be in this category.) One-fourth of these, namely, [b], [m], [v], [z], and [g], do not follow [au] in the monosyllables, and [ŋ] does so only in one uncommon word. Therefore, word-final consonants

are not as generally represented as word-initial consonants, and there are only 15 final sound types.

7. There is no individual initial unit which pairs with each occurring final unit, nor is there any initial sound type which pairs with all occurring final sound types. In fact, there is no individual unit which approximates such a one-to-one relationship. The individual unit most nearly doing so is [r], which pairs with 22 of the 46 final units. Others among the highest frequencies are, in order, [b] and [l] pairing with 20 each; [d] and [g], with 19 and 18, respectively. The [s] type units make up the sound type which pairs with the most different occurring final sound types, 12 out of 15.

8. There is no individual final unit which pairs with each occurring initial unit, nor is there any final sound type which pairs with all the occurring initial sound types. Final [z] is the individual unit which pairs with the most different initial units, 35 out of 47. Other individual final units entering into pairs with relative frequency are [t], [ts], and [d]. The [z] type units make up the final sound type which pairs with the most different occurring initial sound types, in fact, with all occurring initial sound types except the one made up of initial [z] type units. In other words, final [z] type units do not pair with initial [z] type units.

9. If the monosyllables labelled local, dialectal, obsolete, archaic, and rare are removed from the raw data, not only are there fewer specific pairings, but fewer types of units represented in the pairings. There are no pairs with bilabial or labio-dental type final units. There are no pairs with final [ʃ] type units or initial [z] type units. Also, final [k] type units pair only with initial [b], and this in the root and in-

¹⁵ See Betty Jane Wallace, "A Quantitative Analysis of Consonant Clusters in Present-Day English" (Ph.D. dissertation, University of Michigan, 1951), p. 70, for an observation in the same vein.

flected forms of a technical word. All together, 112 of the 398 different pairs found occur only in words so labelled.

10. As is made immediately clear by Table IV, there is a marked preponderance of final alveolar type units in the occurring pairs. No comparable preponderance is observable when the pairs are looked at from the standpoint of the occurring initial units.

With the exception of observation number two, each of the preceding summary statements more or less immediately and urgently raises two closely related general questions: To what

degree is this observation about the monosyllables characteristic of CVC monosyllables in general; to what degree is the situation reported a function of some other factor in the structure of English, for example the frequency of alveolars¹⁶ and the shape of certain inflected forms¹⁷? These and other questions that come to mind point to the need for further research in the field of phonetic context.

¹⁶ See, for example, Rebecca E. Hayden, "The Relative Frequency of Phonemes in General-American English," *Word*, VI (1950), pp. 220 ff.

¹⁷ See Sol Saporta, "Frequency of Consonant Clusters," *Language*, XXXI (1955), p. 27.

THE EFFECT OF CREATIVE ACTIVITIES ON THE ARTICULATION OF CHILDREN WITH SPEECH DISORDERS*

BARBARA M. McINTYRE
University of Pittsburgh

MANY experienced clinicians in the field of speech correction including Van Riper,¹ Backus and Beasley,² and Irwin³ have recognized the therapeutic value of group experience. Some attempts to evaluate the therapeutic effect of group experience on articulation skills have been made. McWilliams⁴ and Ruggieri⁵ found that a planned program of speech activities in a group situation significantly reduced the number of articulation errors in the speech of kindergarten children. Ludwig,⁶ utilizing creative dramatic activities to facilitate auditory training for kindergarten children, found that the experimental group participating in a program of creative dramatics made an improvement in consonant articulation of 52.5 per cent while the children in the control group who did not participate in the creative dramatics showed an improvement of 3.6 per cent.

Since creative dramatics is a group activity involving speech, it appears that

it might meet the needs of the adolescent and pre-adolescent children for therapeutic group experience. Creative dramatics is defined by Ward⁷ as "informal spontaneous drama created by the players themselves." She expresses the opinion that participation in it benefits children by providing the following:

1. Practice in strongly motivated social cooperation.
2. Opportunities to grow in understanding people.
3. Controlled emotional release.
4. Experience in thinking on one's feet and expressing ideas.

These values claimed for creative dramatics parallel, to some extent, some of the goals of speech therapy. Backus and Beasley list as a primary aim in speech therapy "helping each individual to change behavior in inter-personal relationships to the extent that he can function in such relationships with greater relative adequacy in terms of satisfaction and security."⁸

Although there is a lack of research evidence to substantiate claims for the value of creative dramatics in the treatment of disorders in the speech of the school-age child, clinicians at the Speech Clinic of the University of Pittsburgh have for some time utilized creative dramatics as an adjunct to therapy. They have felt that many children benefit from this creative group experience.

*Based in part on Ph.D. dissertation, University of Pittsburgh, 1957, directed by Jack Matthews.

¹ Charles Van Riper, *Speech Correction Principles and Methods* (New York, 1954).

² Ollie Backus and Jane Beasley, *Speech Therapy With Children* (Boston, 1951).

³ Ruth Becky Irwin, *Speech and Hearing Therapy* (New York, 1935).

⁴ Betty Jane McWilliams, "The Effect of Speech Improvement Activities Upon the Consonant Articulation of Kindergarten Children," Master's thesis, University of Pittsburgh, 1950.

⁵ M. K. Ruggieri, "The Effect of a Speech Improvement Project on Consonant Articulation in Four Pittsburgh Kindergartens," Master's thesis, University of Pittsburgh, 1950.

⁶ Charlotte Ludwig, "The Effect of Creative Dramatics Activities Upon the Articulation Skills of Kindergarten Children," Master's thesis, University of Pittsburgh, 1955.

⁷ Winifred Ward, *Playmaking With Children* (New York, 1957), pp. 2-9.

⁸ Backus and Beasley, p. 7.

Because of this there has been an interest at that clinic in evaluating the effect of creative experiences upon the articulation skills of adolescent and pre-adolescent children. It was decided to study the problem experimentally. A six week program of creative activities (creative dramatics, creative music and dance, arts and crafts) for children between the ages of 10 and 14 had been set up, and it was decided to include in this program a number of adolescent and pre-adolescent children who had been receiving therapy for articulation disorders, and to carry out an experiment permitting an estimation of the effect of the program on the articulation skills of these children.

I. EXPERIMENTAL DESIGN

Thirty-two children with articulatory disorders, from the speech therapy program of the Pittsburgh Public Schools, were chosen to participate in the experiment. These children were divided into an experimental and a control group with sixteen subjects in each matched as closely as possible on the basis of chronological age, mental age, socio-economic status, race, and approximate number and type of consonant articulation errors as revealed by an articulation test.

The children in the experimental group participated in a program of creative activities with 185 other children from the Pittsburgh area for three hours each day, five days a week, for a period of six weeks. The members of the control group did not take part in the program of creative activities. At the conclusion of the program the experimental and control groups were retested by the same graduate students who participated in the pre-experimental articulation evaluation, using the same testing instrument.

II. SUBJECTS

During the spring semester of 1955 the names of 331 Pittsburgh school children from speech therapy groups were submitted to the writer by the school therapists. These children ranged in age from ten to fourteen years and all exhibited articulatory defects. Their speech was considered by their therapists to be in need of further attention, although they had been receiving therapy within the school program. The parents of these children were contacted by telephone and by mail, and were asked to bring their children into the University Clinic for articulation testing. The replies indicated that 154 children would be available for the experiment. Of these, 105 kept their appointments and were tested.

Of the 105 children tested, only sixteen were able to attend the program for the full six weeks. The inability of the rest of the children to participate was related to a variety of reasons, such as parents' vacations, distance from the site of the program, illness, transportation difficulties, and lack of parental interest. As previously indicated, the sixteen children in the experimental group were matched as closely as possible according to sex, race, socio-economic level, chronological age, mental age, with sixteen controls. The adequacy of the matching can be seen in the data summary presented in Table I.

It was not possible to match the groups according to the number of consonant articulation errors. The total number of consonant errors for the experimental group was 511, while the total number of errors for the control was 430. The difference between the number of errors in the experimental group and the control group was 81 which yielded a *t* of 2.21, significant at better than the five per cent level of con-

TABLE I
SUMMARY OF STATISTICAL DATA UPON WHICH THE MATCHING OF EXPERIMENTAL AND CONTROL GROUPS WAS BASED.

| | Girls | | Boys | | Combined | |
|---------------------|--------------|---------|--------------|---------|--------------|---------|
| | Experimental | Control | Experimental | Control | Experimental | Control |
| Mean C. A. | 11.1 | 11.1 | 11.3 | 11.2 | 11.3 | 11.2 |
| Mean M. A. | 11.3 | 10.9 | 10.10 | 11.3 | 11.2 | 11.0 |
| Number | | | | | | |
| White | 7 | 7 | 7 | 7 | 14 | 14 |
| Negro | 1 | 1 | 1 | 1 | 2 | 2 |
| Total | 8 | 8 | 8 | 8 | 16 | 16 |
| Articulation Errors | | | | | | |
| Total | 228 | 218 | 283 | 212 | 511 | 430 |
| Mean | 28.50 | 27.25 | 35.37 | 26.50 | 31.94 | 26.87 |
| Standard Deviation | 13.96 | 5.87 | 13.08 | 7.63 | 13.96 | 8.08 |

fidence. A closer matching on consonant errors would have been desirable. However, this would have been difficult. Parents whose children exhibited a great number of errors were anxious to have their children participate in the experimental group, while the parents of the children whose speech exhibited relatively few errors preferred to have their children in the control group, i. e. not participate in the organized six week summer program.

III. PROGRAM OF CREATIVE ACTIVITIES

The program of creative activities was conducted in the open air setting of the park surrounding the Scaife Unit of the Arts and Crafts Centre in Pittsburgh. Here, under the guidance of seven teachers, each holding a Master's degree and having special training in creative arts, 201 children (sixteen of them forming the experimental group) participated in a varied program. Their activities included creative dramatics, creative music and dance, arts and crafts. The program was divided into three sections. Each child took part in each section for one hour a day. The first section was creative dramatics. Here, creative expression through

speech was encouraged. The second section was creative music and dance where bodily coordination was emphasized. The third section was arts and crafts, where expression through manual arts was the objective. No speech correction activities were carried out.

IV. TESTING

A. Pre-Testing of Consonant Articulation.

An oral articulation test was administered to all 105 children who came to the initial testing at the University Clinic. This test was constructed by McWilliams⁹ and was based on the frequency scale compiled by Travis.¹⁰ It represented the relative frequency with which consonant sounds appear in spoken American conversation. Two hundred and five consonant sounds were included in the tests, which were administered by two graduate students well advanced in their master's degree program in speech correction.

⁹ Betty Jane McWilliams, "An Experimental Study of Some Components of Intelligibility of the Speech of Adult Cleft-Palate Patients," Ph.D. dissertation, University of Pittsburgh, 1953.

¹⁰ L. E. Travis, *Speech Pathology* (New York, 1931).

B. Post-testing of Consonant Articulation.

On the final day of the creative activities program the children in both the experimental and the control groups were re-tested by the same testers using the same materials as in the pre-test.

V. ANALYSIS OF THE DATA

A. Total Number of Errors.

From an examination of Table II we see that the experimental group, which started the study with a total of 511 errors and a mean of 31.94, completed the program with a total of 429 errors and a mean of 26.81. This shows a reduction of 82 errors or a mean reduction of 5.13, which yielded a *t* of 3.94 which is significant beyond the one per cent level of confidence.

On the other hand, the control group began the program with 430 errors and a mean of 26.87 errors and on the post-test exhibited 442 errors with a mean of 27.62 errors. The difference in this case yielded a *t* of 1.26 which was not statistically significant.

Therefore, we may conclude that the experimental group which participated in a program of creative activities showed a significant reduction in the number of errors from pre-to post-test while the members of the control group who did not participate in a program

of creative activities did not show a significant change in number of articulation errors.

B. Percentage of Change from Pre-test to Post-test.

Table III shows an analysis of the per cent of change shown by the individual members of the experimental group and the members of the control group from pre- to post-test.

It will be observed that there was a 16.06 per cent reduction of errors made by the experimental subjects and a 2.79 per cent increase in errors among the control subjects. To further evaluate this percentage of change the members of the experimental and control groups were divided into three categories: those showing a decrease in the number of errors, those showing no change, and those showing an increase in the number of errors from pre- to post-test.

Table IV shows that a significantly greater percentage of children in the experimental group decreased the number of errors from pre-test to post-test than did the children in the control group. Furthermore, a greater percentage of children in the control group retained the same number of errors from pre- to post-test than did the members of the experimental group. There was no significant difference between the

TABLE II
COMPARISON OF PRE-TEST AND POST-TEST ERROR MEANS FOR EXPERIMENTAL AND CONTROL GROUPS.

| | Experimental (N-16) | | | Control (N-16) | | | | | | |
|------------------------------|---------------------|----------|-----------|----------------|--------------------|---------------|----------|-----------|-------|--------------------|
| | No. of Errors | Pre-test | Post-test | Diff. | Per Cent of Change | No. of Errors | Pre-test | Post-test | Diff. | Per Cent of Change |
| Total | 511 | 429 | 82 | 16.06 | | 430 | 442 | 12 | 2.79 | |
| Mean | 31.94 | 26.81 | 5.13 | 1.00 | | 26.87 | 27.62 | .75 | .17 | |
| Stand. Dev. | 13.96 | 16.14 | 2.18 | | | 8.08 | 9.59 | 1.51 | | |
| <i>t</i> Ratio between Means | 3.94* | | | | | 1.26 | | | | |

*Significant at 1% level

TABLE III

TOTAL NUMBER OF ERRORS AND THE PER CENT OF CHANGE SHOWN BY INDIVIDUAL MEMBERS OF THE EXPERIMENTAL AND CONTROL GROUPS FROM PRE- TO POST-TEST.

| Experimental (N-16) | | | | Control (N-16) | | | | | |
|---------------------|----------|-----------|-------|--------------------|---------------|----------|-----------|--------|--------------------|
| No. of Errors | Pre-test | Post-test | Diff. | Per Cent of Change | No. of Errors | Pre-test | Post-test | Diff. | Per Cent of Change |
| 16 | 11 | 5 | 5 | 31.25 | 13 | 13 | 0 | 0.00 | 00.00 |
| 33 | 19 | 14 | 5 | 42.42 | 21 | 21 | 0 | 0.00 | 00.00 |
| 54 | 58 | —4 | —4 | —7.40 | 32 | 34 | —2 | —6.25 | |
| 52 | 45 | 7 | 7 | 13.46 | 36 | 47 | —11 | —30.55 | |
| 45 | 57 | —12 | —12 | —26.66 | 30 | 27 | 3 | 10.00 | |
| 14 | 6 | 8 | 2 | 57.14 | 30 | 30 | 0 | 0.00 | |
| 19 | 27 | —8 | —8 | —42.10 | 20 | 21 | —1 | —5.00 | |
| 38 | 21 | 17 | 4 | 44.73 | 48 | 49 | —1 | —2.04 | |
| 16 | 3 | 13 | 10 | 81.25 | 23 | 24 | —1 | —4.34 | |
| 38 | 29 | 9 | 20 | 29.68 | 25 | 25 | 0 | 0.00 | |
| 44 | 27 | 17 | 10 | 38.63 | 34 | 34 | 0 | 0.00 | |
| 46 | 38 | 8 | 30 | 17.39 | 32 | 31 | 1 | 3.12 | |
| 40 | 36 | 4 | 32 | 10.00 | 17 | 17 | 0 | 0.00 | |
| 18 | 18 | 0 | 18 | 00.00 | 17 | 18 | —1 | —5.88 | |
| 14 | 9 | 5 | 4 | 35.71 | 27 | 27 | 0 | 0.00 | |
| 24 | 25 | —1 | —1 | —4.17 | 25 | 24 | 1 | 4.00 | |
| 511 | 429 | 82 | 327 | 16.06 | 430 | 442 | 12 | —2.79 | |

— indicates increase in error.

TABLE IV

COMPARISON BETWEEN THE PERCENTAGES OF CHILDREN IN THE EXPERIMENTAL AND CONTROL GROUPS SHOWING A DECREASE IN ERROR, NO CHANGE, AND AN INCREASE IN ERROR FROM PRE- TO POST-TEST, AND THE *t* RATIO OBTAINED.

| | Decrease in Error No. | Decrease in Error Per Cent | | No Change No. | No Change Per Cent | | Increase in Error No. | Increase in Error Per Cent |
|----------------|--------------------------|-------------------------------|--|------------------|-----------------------|--|--------------------------|-------------------------------|
| Experimental | 11 | 68.75 | | 1 | 6.25 | | 4 | 25.00 |
| Control | 3 | 18.75 | | 7 | 43.75 | | 6 | 37.50 |
| Difference | 8 | 50.00 | | 6 | 37.50 | | 2 | 12.50 |
| <i>t</i> Ratio | 3.3** | | | 2.72* | | | .77 | |

*Significant at 5% level.

**Significant at 1% level.

control group and the experimental group in percentage showing an increase in error from pre-test to post-test.

C. Rating Scale Evaluation of Errors.

During the evaluation of the results of this study, it appeared that there were movements along the error continuum from omission through substitution and distortion to correctly articulated sounds. For example, when a child began the study with an omission but completed the study with a distortion of the same sound, we might assume that progress had been made.

Support for this theory can be found in the work of Milisen¹¹ and his associates. Therefore, a three point scale was devised which provided a rating of one for a distortion, a rating of two for a substitution, and a rating of three for an omission. According to this scale, known as the P-Scale, a low score is indicative of relatively more mature articulation, i. e., articulation closer to normal, than is a high score. Table V

¹¹ Robert Milisen and Associates, "The Disorder of Articulation: A Systematic Clinical and Experimental Approach," *Journal of Speech and Hearing Disorders*, Monograph Supplement No. 4, (1954).

TABLE V

COMPARISON OF P-SCALE SCORES OF EXPERIMENTAL AND CONTROL GROUPS AND THE *t* RATIO OBTAINED BETWEEN EXPERIMENTAL PRE- AND POST-TESTS AND CONTROL PRE- AND POST-TESTS.

| | P-Scale Score Total | | | Mean | | | <i>t</i> Ratio Between Pre- and Post-Tests | | |
|--------------|---------------------|------|-------|-------|-------|-------|--|--|--|
| | Pre | Post | Diff. | Pre | Post | Diff. | | | |
| Experimental | 783 | 690 | 93 | 48.93 | 43.12 | 5.81 | 2.48* | | |
| Control | 538 | 561 | 23 | 33.50 | 35.06 | 1.56 | .83 | | |

*Significant at 5% level.

shows a comparison of P-Scale scores earned by the experimental and control group on pre and post-tests.

It will be observed that the experimental group made a significant reduction in the P-Scale score from pre to post-test and thus, if the Milisen theory is correct, made progress along the continuum toward the goal of correctly articulated consonant sounds. The control group, on the other hand, did not show such improvement.

D. Evaluation of Errors of Boys and Girls.

An analysis in the difference, if any, of the effect of creative activities upon the articulation skills of boys compared with girls was undertaken, the results of which are given in Table VI. It will be observed that among the experimental subjects the girls made a significantly greater reduction of errors than did the boys.

E. Types of Error.

Roe and Milisen¹² suggest that, at the adolescent age level the primary type of articulation error is likely to be the distorted consonant sound. Table VII presents a summary of data collected in the present study concerning the number of distortions, substitutions, and omissions on pre- and post-test in the experimental and the control groups.

The results there given indicate that the only significant change in type of error from pre-test to post-test was exhibited by the experimental group in the reduction of distortions.

F. Frequency of Sound Errors.

Examination of the data collected in the present study showed that the most frequent errors were in the produc-

¹² Vivian Roe and Robert Milisen, "The Effect of Maturation Upon Defective Articulation in Elementary Grades," *Journal of Speech Disorders*, 7 (1942), 44-45.

TABLE VI

COMPARISON OF THE NUMBER, MEAN AND DIFFERENCE IN ERRORS FOR BOYS AND GIRLS AND THE *t* RATIOS OBTAINED BETWEEN BOYS' PRE- AND POST-TESTS AND GIRLS' PRE- AND POST-TESTS IN BOTH EXPERIMENTAL AND CONTROL GROUPS.

| | Experimental (N-16) | | | | | | Control (N-16) | | | | | | <i>t</i> Ratio between Pre and Post | |
|--------------|---------------------|-----|------|-------------|-----|------|----------------|-------|-------|-------------|------|-------|-------------------------------------|-------|
| | Boys (N-8) | | | Girls (N-8) | | | Boys (N-8) | | | Girls (N-8) | | | | |
| | No. of Errors | Pre | Post | Diff. | Pre | Post | Mean of Errors | Post | Diff. | Pre | Post | Diff. | | |
| | B | G | B | G | B | G | B | G | B | G | B | G | | |
| Experimental | 283 | 228 | 250 | 170 | 33 | 49 | 35.37 | 28.50 | 31.25 | 22.87 | 4.12 | 6.13 | 1.76 | 3.24* |
| Control | 212 | 218 | 219 | 223 | 7 | 5 | 26.50 | 27.25 | 27.37 | 27.87 | .87 | .62 | .62 | .20 |

*Significant at 5% level.

TABLE VII

COMPARISON BETWEEN NUMBER OF CONSONANT DISTORTIONS, SUBSTITUTIONS, AND OMISSIONS FROM PRE-TEST TO POST-TEST IN EXPERIMENTAL AND CONTROL GROUPS.

| Type of Error | Experimental | | | | <i>t</i> Ratio | Control | | | | |
|---------------|-------------------|--------------------|-----------------|------------------|----------------|-------------------|--------------------|-----------------|------------------|----------------|
| | No. of Errors Pre | No. of Errors Post | % of Errors Pre | % of Errors Post | | No. of Errors Pre | No. of Errors Post | % of Errors Pre | % of Errors Post | <i>t</i> Ratio |
| Dis. | 336 | 258 | 65.76 | 60.14 | 3.03** | 350 | 346 | 81.39 | 78.51 | .27 |
| Sub. | 78 | 81 | 15.26 | 18.88 | .01 | 52 | 73 | 12.09 | 16.51 | .12 |
| Om. | 97 | 90 | 18.98 | 20.97 | .31 | 28 | 23 | 6.52 | 5.20 | .59 |

**Significant at 1% level.

tion of [s] and [z] sounds. The [s] sound was misarticulated 71.52 per cent of the times it was attempted on the pre-test in both the experimental and the control groups. On the post-test it was misarticulated 50.65 per cent of the times it was attempted by the experimental group and 74.67 per cent of the times by the control group. The [z] sound was misarticulated 62.50 per cent of the times it was attempted on the pre-test by the experimental group and 59.37 per cent by the control group. Examination of the post-test data showed [z] to be in error 44.37 per cent of the times it was attempted in the experimental group and 60.00 per cent in the control group. The next most frequently misarticulated sounds were [r], [ʃ], [θ]. However, errors in these sounds appeared much less frequently than did errors in [s] and [z].

VI. CONCLUSIONS

1. The experimental group, which participated in the program of creative

activities, made a significant reduction in the number of consonant articulation errors.

2. A greater percentage of the children in the experimental group than in the control group showed improvement in their consonant articulation skills.

3. The experimental group evidenced significantly greater progress along the error continuum toward correctly articulated sounds than did the control group.

4. Girls in the experimental group made a significant reduction in the number of consonant articulation errors from pre- to post-test, while the boys in the same group did not.

5. Distortion was the only type of error that exhibited significant change from pre- to post-test.

6. The greatest reduction in errors on individual sounds was found to have occurred in the experimental group in the production of [s] and [z].

AN
C
with
used
the
desc
spee
emp
crite
spee
appe
Wher
twee
ficult
preta
both
addre
pre-s
al sta
The
the
rheto
speak
T
and
effor
but
the
and
R
The
stud
natu
mar
*Bas
South
C. D
1
Criti
2

AN EXPERIMENTAL STUDY OF AUDIENCE RECOGNITION OF EMOTIONAL AND INTELLECTUAL APPEALS IN PERSUASION*

RANDALL C. RUECHELLE
Colorado State University

I. INTRODUCTION

CONTINUAL questions arise in the study of persuasion having to do with the types of persuasive appeals used by a speaker and whether or how the appeals might be classified or described. From the point of view of speech criticism, Thonssen and Baird¹ emphasize the necessity for finding criteria for judgment of the effect of a speech. In a discussion of persuasive appeals they state:

When we come to analyze the differences between the two types of appeal, we run into difficulties—difficulties both of definition and interpretation. Common sense tells us, of course, that both thought and feeling function in public address; that is, the notion of communication pre-supposes both an ideational and an emotional state in the speaker. . . .

The problem affects the listeners no less than the speaker. They are respondents in the rhetorical process; their reaction is, to the speaker, the all-important consideration.²

The substance of what is emotional and what is intellectual in a persuasive effort may well be difficult to determine, but any study of persuasion presumes the existence of thought and feeling, and often presumes their effect.

Review of Previous Investigations. There have been various experimental studies which have touched upon the nature of persuasive appeals; the primary concern has been the attempt to

*Based on Ph.D. dissertation, University of Southern California, 1953, directed by Milton C. Dickens.

¹ Lester Thonssen and A. Craig Baird, *Speech Criticism* (New York, 1948).

² Thonssen and Baird, pp. 357-358.

determine the relative effectiveness of types of appeals. Collins³ made a study employing four speeches written by himself comprising various combinations of "emotional" and "logical" arguments. After analyzing the reactions of auditors, he concluded that the speech which had intermingled appeals was most effective. Millson⁴ undertook a study of relative effectiveness of logic and emotion. He concluded that "emotional" speaking had greater influence. A similar study by Menefee and Granneberg⁵ arrived at a similar conclusion.

Knower's⁶ study of audience attitude measurement employed speeches "predominantly factual or logical," and speeches "predominantly emotional." Recognizing the difficulties of such a procedure, he submitted the texts to ten teachers of speech for agreement as to their type. The portion of his conclusions which related to the nature of the appeals used in the speeches stated that "logical and persuasive speeches were equally effective in producing changes of attitude."⁷

³ G. Rowland Collins, "The Relative Effectiveness of the Condensed and Extended Motive Appeal," *Quarterly Journal of Speech*, X (1924), 221-230.

⁴ W. A. D. Millson, "Problems in Measuring Audience Reaction," *Quarterly Journal of Speech*, XVIII (1932), 621-637.

⁵ S. C. Menefee and A. G. Granneberg, "Propaganda and Opinions on Foreign Policy," *Journal of Social Psychology*, XI (1940), 393-404.

⁶ Franklin H. Knower, "Experimental Studies of Changes in Attitudes: I. A Study of the Effect of Oral Argument on Changes of Attitude," *Journal of Social Psychology*, VI (1935), 315-347.

⁷ Knower, p. 342.

The source for construction or compilation of the appeals used in these studies was the experimenter himself. Monroe commented on investigations of this type:

... These studies were concerned with only a few of the many variables involved in the speaker's performance. Further, the attempt was made in each case to *define in advance* the variable to be studied and then to isolate it experimentally. While this is in most cases a commendable experimental technique, with the speaker-audience phenomenon it may lead to doubtful results because we do not yet know what are the independent variables of the speaker's performance as the *auditor observes them*. There is no certainty, for example, that logical argument and emotional appeal are distinct from one another, or even that they exist as unitary variables.⁸

Statement of Problem. This investigation sought to examine the patterns of audience recognition of persuasive appeals. It held to the hypothesis that if persuasive appeals can be categorically classified as emotional or intellectual, the source for such classification might be found in recognition by the observers of these appeals. Therefore, if the patterns of recognition were such that a definite distinction were discernible, a basis for classification could be established. If no such distinction were discernible, a classification of appeals based on observer recognition has no meaning.

For purposes of this investigation, persuasive appeals were considered in terms of degree of recognizable amounts of emotional or intellectual content in a persuasive effort. It was realized that an appeal might well be successful without the observers being fully aware of how much or how little they had actually been influenced in behavior or belief, yet it was further realized that any

⁸ A. H. Monroe, "The Measurement and Analysis of Audience Reaction to Student Speakers—Studies in Attitude Changes," Bulletin of Purdue University, 38, Studies in Higher Education, XXXII (1937), pp. 8-9.

attempt to measure objectively the immediate reactions of the observers had to depend upon the recognition by the observers. It is upon this basis that classifications of appeals are frequently made.

The problem under investigation was whether persuasive appeals could be classified categorically as to amounts of emotional and intellectual content. Specific questions were formulated in an attempt to find answers to the problem.

1. In the pattern of reactions of auditors, viewers, or readers of persuasive materials, could the following criteria be determined: (a) Did reactions show a dichotomous pattern: was material which was rated as more emotional correspondingly rated as less intellectual, and vice versa? (b) Did ratings of persuasive materials correspond to the speaker's intent? (c) Did ratings of persuasive materials show uniform agreement as to type and degree of appeal? (d) What were the proportional bases for judgment expressed by the ratings: content, delivery or wording, general impression, or unknown?

2. Did various factors influence the general patterns of reactions as expressed by the ratings? (a) Was there a sex difference in ratings? (b) Were ratings influenced by agreement or disagreement with the speakers' points of view? (c) Did the general distribution of ratings differ according to the basis of judgment?

3. Was there a marked similarity or dissimilarity in reaction patterns of various groups receiving the material through different media? (a) How did groups react who heard the recorded speeches? (b) How did groups react who both heard and saw the speeches? (c) How did raters who were inexperienced in speech react to the persuasive materials in print? (d) How did "expert"

raters trained and experienced in speech react to the persuasive materials in print?

II. EXPERIMENTAL PROCEDURES

In the creation of the materials, it was necessary to combine into one medium a system of keeping those materials constant and uniform and also to allow them to be prepared for presentation in different forms in the separate testings. In the first testing the materials were to be presented audibly. In the second testing the materials were to be presented visually and audibly, and in the third testing the materials were to be presented in manuscript form. The principal medium selected for a permanent and, at the same time, most life-like recording and preservation was sound motion picture film.

Preparation of Materials. The materials constituted short speeches each embodying a single persuasive appeal. The question arose as to how long a particular speech should be in order to embody a recognizable appeal. Persuasive speeches employed in previous studies have varied in length from one minute to fifteen minutes or longer. Theoretically, a word or two might constitute an appeal as might a speech of indeterminable length. The major consideration was to obtain a variety of appeals on a variety of subjects by a variety of speakers and to put them into form for a practical rating procedure for groups of judges. It was decided to obtain speeches of uniform length and that a time limit of two minutes be maintained. Such a time limit would fulfill practical considerations in the experiment and provide an adequate opportunity for a speaker to make his intent known.

Twenty-one male speakers were selected to make the persuasive speeches.

These speakers were mature and experienced performers. The group included four speech instructors, five graduate students in public address, five upper division speech majors, four senior debaters, and three ministers. The ages ranged from twenty to fifty-five, the median age being twenty-eight. The length of speaking experience ranged from three to twenty-five years, the median being nine and one-half years.

A list of suggested topics was provided for the assistance of the speakers, although they were given full option of rephrasing any topic of their own choosing and of speaking on either side of a topic as it was stated in the list. The list of suggested topics was drawn from a more inclusive list submitted to twenty-six faculty members and graduate students of speech at the University of Southern California. The original list was graded on the basis of significance of subject, adaptability to development of a concrete view in a two-minute limit, and general interest. Mean preferential ranks were computed and used as the basis for selection of the twelve suggested topics for the speakers.

A single day was scheduled for the photographing of the speeches. After completing his speech, each speaker in turn was conducted from the studio and handed a rating form on which he was asked to state whether he intended to use primarily an emotional or intellectual appeal. He was further instructed to indicate the relative degree of emotional and intellectual appeal he used. The intent of the speakers, then, was considered not as their deliberate effort to employ a designated degree of emotionality or intellectuality, but rather the self-ratings of their performance immediately after the persuasive speech was made.

The experimenter edited and spliced

the films and included twenty seconds running time of blank leader film after each speaker's performance. The total running time of the films was fifty minutes.

The second step in preparation of materials was to obtain sound recordings of the speeches. The sound track of the films was transcribed onto reels of magnetic recording tape. The same sequence of the speeches and the same twenty-second intervals between the speeches was maintained.

The third step in preparation of materials was the transcription of the exact words of the speakers from the tape recordings into printed texts of the speeches.

Rating Procedures. The method of determining audience reaction to the given material was through use of rating scales. An open-end linear scale on which five categories were stated was considered most advisable to correspond with the previous assumption that amount implies a quantitative degree and the measurement of that quantitative degree would provide possible justification for categorization of the materials to be tested. The ratings or judgments would, then, be represented by a method of successive intervals. A limited number of categories was stated and arranged in rank order. The rater was to evaluate the amount of the appeal by placing a check mark in one of these categories. No presumption of equality of distances between neighboring categories was made. The scales that appeared on the rating sheets took this form:

| 1 () | 2 () | 3 () | 4 () | 5 () |
|----------------------------|--------------------------------|--------------------------|--------------------------------|----------------------------------|
| Almost no emotional appeal | Below average emotional appeal | Average emotional appeal | Above average emotional appeal | Almost entirely emotional appeal |

An identical scale was used to determine the amount of intellectual appeal. In that scale, the word intellectual was substituted for the word emotional.

In order to determine the general attitude of the auditors toward the subjects on which the speakers were to speak, a pre-test questionnaire was used. It consisted of a list of the topics used by the speakers. Raters were asked to indicate their attitude toward the topic by checking yes, no, or undecided for each of the topics.

In order to determine the basis of judgment of the amount of emotional and intellectual appeals, a check list system was employed. Raters were asked to check one of four items:

I based my judgment of the amounts of emotional and intellectual appeals in the speech mainly on *one* of the following:

1. *What the speaker said:* _____ (Content of the speech)
2. *How the speaker spoke:* _____ (Delivery of the speech)
3. *General impression of the speaker:* _____ (Content and delivery combined)
4. *(I am) Undecided:* _____

A rewording of this check list was necessary for those rating the manuscripts of the speeches. In the third test the following terms were included: content, wording, general impression, cannot distinguish between content and wording.

Subjects. Undergraduate students were used as subjects in the first two testings. The groups were recruited from the beginning speech course taught in the Department of Speech at the University of Southern California. The first test consisted of playing the tape record-

ings of the speeches. This was accomplished in six separate sessions before 151 students, each session averaging approximately twenty-five students. Ratings were obtained from 114 male students and 37 female students.

The second test consisted of projection of the sound motion pictures. This was accomplished in four separate sessions. Fifty beginning speech students served as subjects. Forty were male students, ten were female. Each of the four sessions averaged twelve to thirteen students.

The ages of the subjects in the first two tests ranged from seventeen to fifty-one, the median age being 21.7 years. All four college years were included. The mean years in college for all subjects was 2.3 years.

The third test involved the rating of the speech manuscripts. Two separate groups of thirty raters each were employed. The first group consisted of thirty adult raters chosen at random. The only qualification for selection was that none of the raters was to have any training or experience in the study or teaching of speech. This group was designated as "lay" raters. The second group consisted of thirty individuals who had experience in teaching courses in speech at either or both high school and college levels. These raters were characterized as "expert" raters. They were recruited from speech staffs of six colleges and universities in the Los Angeles area. Their experience in teaching courses in speech ranged from two to thirty-three years, with a median of 9.6 years. Fourteen held doctorates; sixteen held master degrees. All had either a major or minor in rhetoric and public address in their graduate work.

Testing Procedures. In the first and second tests, conditions were controlled so as to insure equal representations of

materials to the various groups involved. All instructions were presented in printed form. The rating procedure involving use of the manuscripts was not controlled. Raters were asked to indicate the difficulty of the test. Twenty-four per cent indicated that it was "hard."

III. DATA AND INTERPRETATION

Dichotomic Pattern of Ratings. The question was posed whether certain criteria could be determined from the reactions of raters. Basic to the ability of raters to decide whether an appeal was more or less emotional or intellectual would be the dichotomic nature of their ratings. In terms of the hypothesis, the categories should be mutually exclusive. That is, if an appeal were to be classified as more emotional, it would be presumed to be correspondingly less intellectual, and vice versa. If such a pattern could be determined from the nature of the ratings, the hypothesis would be sustained. A dichotomic pattern would be evident if the distribution of ratings showed a linear regression. A relationship shown by a substantial negative correlation would indicate ability of raters to classify categorically the persuasive appeals.

To determine the dichotomic patterns of ratings, Pearson product-moment coefficients of correlation were employed using the basic formula,⁹ and all obtained correlations for individual speeches were corrected for coarse grouping by use of the following formula:¹⁰

$$r_e = \frac{r}{\sqrt{C_x C_y}}$$

⁹ J. P. Guilford, *Fundamental Statistics in Psychology and Education* (New York, 1950), pp. 157 ff.

¹⁰ Charles C. Peters and Walter R. Van Voorhis, *Statistical Procedures and Their Mathematical Bases* (New York, 1940), pp. 398 ff.

TABLE I
CORRELATIONS OF EMOTIONAL WITH INTELLECTUAL RATINGS
FOR INDICATION OF DEGREE OF AGREEMENT
BY DICHOTOMIC NATURE OF THE RATINGS.

| Speech No. | Coefficients of Correlation (Corrected r 's) | | |
|---------------------------------|--|---------|-------------------|
| | Test I | Test II | (Lay) Test III |
| | N = 151 | N = 50 | N = 30 |
| 1 | -.263** | -.183 | -.192 |
| 2 | -.556** | -.030 | .102 |
| 3 | -.310** | .402** | .558** |
| 4 | -.186* | -.133 | .213 |
| 5 | .204** | .293* | .123 |
| 6 | .235** | .398** | -.411* |
| 7 | -.015 | -.006 | -.026 |
| 8 | .005 | -.363* | -.171 |
| 9 | -.211** | -.383** | .266 |
| 10 | .082 | .012 | .121 |
| 11 | .210** | .075 | -.345* |
| 12 | -.086 | -.071 | -.036 |
| 13 | -.254** | .204 | .156 |
| 14 | -.195* | -.543** | -.095 |
| 15 | -.043 | -.376** | .213 |
| 16 | .043 | -.060 | -.308* |
| 17 | .072 | -.091 | .279 |
| 18 | -.040 | -.784** | .015 |
| 19 | .215** | .470** | .324* |
| 20 | -.018 | .064 | .278 |
| 21 | -.146 | .350* | -.029 |
| Average (by z transformation) | -.034 | -.009 | .052 |

*Significant at 5% level.

**Significant at 1% level.

Average correlations for all speeches were obtained by use of Fisher's z transformation.¹¹ Results are reported in Table I.

It will be observed that the first two tests resulted in a rating pattern that showed almost no relationship from which to assume or predict a dichotomy. The categories were not mutually exclusive according to the judgments of the raters. The ratings of lay raters in the third test showed an actual positive relationship of small degree. The ratings of experts showed a larger negative relationship reflecting a higher degree of discrimination but one which was neither significant for prediction of categorical division nor as an index to support assumption of mutually exclusive categories.

No substantial support of the hypoth-

esis that classification of appeals could be determined from recognition by raters was to be found. The lack of relationship might be even more pertinently illustrated by application of the coefficient of alienation formula:¹²

$$k = \sqrt{1 - r^2}$$

With intellectual ratings represented by x and emotional ratings represented by y , k_{xy} for Group I was equal to a figure in excess of .999. This would indicate a percentage reduction in errors of prediction of y from x of less than .1 and a percentage of variance accounted for of approximately 0.0%. A similar interpretation was apparent for Group II. That is, the percentage reduction in errors of y from x was less than .1, k_{xy} being in excess of .999. The percentage of variance accounted for was also less than 1%. A comparison between the

¹¹ Guilford, pp. 609-610.

¹² Guilford, p. 408.

TABLE II
COMPARISON OF COEFFICIENTS OF CORRELATION OF RATERS
IN TEST III. SIGNIFICANCE IN TERMS OF k

| Groups | r_{xy} | k_{xy} | 100 $(1-k_{xy})$ % reduction in errors of prediction of y from x | 100 r^2_{xy} % of variance accounted for |
|------------------|----------|----------|--|--|
| Lay Raters | .052 | .999 | .1 | 0.0 |
| Expert Raters | -.286 | .957 | 4.3 | 8.38 |

groups in Test III might be illustrated by tabular representation of k coefficients of alienation. Such a comparison is reported in Table II.

Speakers' Intent. The intent of the speakers was the self-rating given by each speaker to his own performance. It remained to be seen whether the ratings derived from the tests agreed with the speakers' self-ratings on both scales. The results of corrected correlations of ratings with the speakers' self-ratings are reported in Table III.

two out of three times within the limits of +.134 and +.294. The low positive relationship would have little meaning. In general, the assumption that agreement of ratings with a speaker's intent would provide a basis for classifying appeals had no substantial support.

Factors Influencing Ratings. It seemed appropriate to determine what factors might influence judgments of raters and to compare consistency of ratings for each scale considered separately. Such examination might help clarify the pat-

TABLE III
CORRELATIONS (CORRECTED r 's) OF RATINGS WITH SPEAKERS'
SELF-RATINGS TO INDICATE AGREEMENT WITH
SPEAKERS' INTENT

| Scales | Test I N = 151 | Test II N = 50 | (Lay) Test III N = 30 | (Experts) Test III N = 30 |
|------------------------|-------------------|-------------------|-----------------------------|---------------------------------|
| Emotional Scales | .214** | .208 | .175 | .276 |
| Intellectual Scales | .062 | .118 | -.003 | .082 |

**Significant at 1% level.

It was noted that the correlation for emotional ratings was markedly higher than the correlation for intellectual ratings. However, the degree of agreement was low for each group on each scale. Only for ratings given emotional content in the first test (recordings) could it be predicted that raters of emotional content of recorded persuasive speeches would agree with speakers' self-ratings

terms of reactions. Ratings in the first test employing 114 male subjects and 37 female subjects were used to ascertain whether a significant sex difference was apparent. The comparison of ratings by sex is reported in Table IV. The significance of differences between proportional ratings by categories on both emotional and intellectual scales showed t ratios of no significance. It was evident

TABLE IV
COMPARISON OF RATINGS BY SEX IN TEST I

| Category | EMOTIONAL SCALE | | | |
|----------|------------------|--------------------|------------|-----------------|
| | <i>f</i> Male | <i>f</i> Female | σd | <i>t</i> ratios |
| 1 | 201 | 66 | .0015 | .079 |
| 2 | 485 | 147 | .0163 | .822 |
| 3 | 799 | 259 | .0195 | .026 |
| 4 | 640 | 219 | .0185 | .789 |
| 5 | 269 | 86 | .0130 | .131 |

| Category | INTELLECTUAL SCALE | | | |
|----------|--------------------|--------------------|------------|-----------------|
| | <i>f</i> Male | <i>f</i> Female | σd | <i>t</i> ratios |
| 1 | 209 | 69 | .0117 | .128 |
| 2 | 527 | 188 | .0175 | 1.251 |
| 3 | 927 | 306 | .0202 | .327 |
| 4 | 595 | 177 | .0174 | 1.190 |
| 5 | 136 | 37 | .0097 | 1.022 |

TABLE V
PROPORTIONAL BASIS FOR JUDGMENT EXPRESSED
BY THE RATINGS

| Basis | Test I N = 151 | Test II N = 50 | (Lay) Test III N = 30 | (Experts) Test III N = 30 |
|------------------------------|-------------------|-------------------|-----------------------------|---------------------------------|
| Content | .363 | .341 | .390 | .390 |
| Delivery (Wording, Test III) | .142 | .159 | .180 | .146 |
| General Impression | .411 | .467 | .390 | .450 |
| Undecided | .084 | .033 | .040 | .014 |

that sex difference had no significant bearing on distribution of ratings or judgment of the persuasive content.

Ratingers tended to base judgments mainly on general impression, less on content, and still less on delivery or wording as seen in Table V.

The agreement, disagreement or indecision of raters concerning the topics on which the speakers spoke was considered a potential influence on the raters' judgments. In order to determine whether such an influence was evident in the groups tested, a comparison of rating distributions by categories on each scale was made with the three categories of agreement. The purpose was to determine whether the same population could be assumed for the categories of agreement.

Chi-square was employed as a test of the null hypothesis that agreement or disagreement with the speakers' points of view would have no bearing on the rating of the emotionality or intellectuality of what he said.¹³ The resultant χ^2 's are reported in Table VI.

Ratings of emotional content were relatively constant throughout. Influence of agreement or disagreement could not generally be predicted. However, ratings of intellectual content were relatively predictable according to disagreement with the speakers' points of view. The tendency revealed in the distribution of ratings was for the rater to mark the intellectual content lower if he disagreed with the speaker, higher if he agreed with the speaker.

¹³ Guilford, p. 276.

TABLE VI

COMPARISON OF RATINGS WITH AGREEMENT, DISAGREEMENT, OR INDECISION OF RATERS CONCERNING SPEAKERS' POINTS OF VIEW

| Scales | χ^2 's | | | Test III (Experts) |
|---------------------|-------------|----------|-------------------|-----------------------|
| | Test I | Test II | Test III (Lay) | |
| Emotional Scales | 10.63 | 11.69 | 23.79* | 6.25 |
| Intellectual Scales | 29.27** | 14.21*** | 19.15* | 23.12** |

*Significant at 5% level.

**Significant at 1% level.

***Significant between 10% and 5% levels.

TABLES VII

COMPARISON OF RATINGS ACCORDING TO BASIS OF RATING (CONTENT, DELIVERY OR WORDING, GENERAL IMPRESSION, UNDECIDED)

| Scales | χ^2 's | | | Test III (Experts) |
|---------------------|-------------|---------|-------------------|-----------------------|
| | Test I | Test II | Test III (Lay) | |
| Emotional Scales | 73.94** | 44.90** | 20.93** | 8.35 |
| Intellectual Scales | 47.86** | 72.61** | 33.01** | 36.76** |

**Significant at 1% level.

The question remained whether the distribution of ratings varied significantly according to the principal bases for judgment. Here again, the purpose was to determine if the same population could be assumed from a comparative analysis.

A chi-square test was made of the null hypothesis that the basis upon which judgments were made would have no bearing on the ratings of the emotional or intellectual content of the speeches. The resultant χ^2 's are reported in Table VII.

It was seen that ratings were, for the most part, influenced by the basis of judgment used in the rating. That is, different patterns were generally revealed for those rating primarily on content, or delivery, or wording, or general impression, or something unknown.

IV. CONCLUSIONS

1. The series of unclassified persuasive materials presented in the tests could *not* be dichotomized or classified by the observers as emotional or intellectual in content.
2. There was no apparent consistency in classifying like materials within any one group tested in a single test, or between groups in the separate tests.
3. Individual auditors and readers reacted differently to like materials under like circumstances.
4. There was generally only insignificant agreement of the raters with the speakers' self-ratings.
5. Raters in each group in each test tended to base judgments mainly on general impression, less on content, still less on delivery or wording; very few were undecided.

6. Over three-fourths of the raters did not consider the tests to be difficult.
7. There was no apparent consistency in influence of ratings according to disagreement with the speakers' points of view, although a tendency to rate lower on intellectual content if the raters were in disagreement was noted.
8. There was no apparent consistency in evident differences of rating patterns according to the main bases of judgment.
9. Use of the different media of presentation to different groups made little difference in the rating patterns.
10. "Expert" raters were little better able to determine the nature of the appeals than "lay" raters chosen at random. Neither group had significant agreement.
11. It appeared that the hypothesis of the study was negated since the source for classification of materials could not be found in the recognition of appeals by auditors or readers. Therefore, the assumption of a clear-cut classification of emotional and intellectual appeals in persuasion had no discernible basis insofar as the examination undertaken in this study was concerned.

HERMOGENES ON "STOCK ISSUES" IN DELIBERATIVE SPEAKING

RAY NADEAU
University of Illinois

IN glancing through the indices of current texts dealing, in whole or in part, with preparation for deliberative speaking, one will usually find page references to "stock issues." No formal treatment of debate, for example, is considered complete without a set of such issues in one guise or another; the number may vary from two to seven or more, but it seems that the debater cannot get along without these standard approaches to the analysis of a proposition.

Dieter defined stasis, for which the popular current term is "issue," as ". . . the rest, pause, halt, or standing still which inevitably occurs between opposite as well as between contrary 'moves,' or motions."¹ As applied to rhetoric, a stasis or issue is a point in controversy which acts as a focus or center for opposing contentions. "Is international control of the Suez necessary?" The foregoing is a question indicating a position of balance or "rest" between opposing forces—proponents contend that the proposed action is necessary; opponents contend that it is not. When opposing points of view are centered on such a key question, that question is designated by theorists as a main issue. If the opposing views are centered on a part or subdivision of a key question, the resulting issue is described as being secondary or subordinate. In other words, stases or issues come into being as a result of conflict over key questions or over secondary

points arising in the course of discussion of those questions. As Quintilian says, it is not until the parties cease to agree that any question (i.e. issue) arises (vii. 1. 6.)

What is a *stock* issue? In his preparation for deliberative speaking (as well as for forensic and epideictic occasions), man long ago discovered that certain questions occurred so frequently that an orderly listing of them provided a convenient pattern for analysis as well as a guide to the proper phrasing of specific issues related directly to a particular proposal. For example, "Is the proposed action necessary?" Here we have a stock issue that could very quickly become a specific main issue in the consideration of most deliberative questions, including the one on the Suez problem in the paragraph above. A stock issue, then, is a *possible* issue, general in its phrasing, which may or may not become an actual and specific, main or secondary issue in the discussion of a definite proposition.

However popular stock issues may be among debaters, legislators, and others concerned primarily with what ought to be done now or in the future, these issues are not the contribution of any modern theorist—the ancients used them and established patterns that have been traditional for centuries. It is the purpose of this paper (1) to sketch the background of the development of stock issues in deliberative speeches and (2) to present the first modern language translation of what Hermogenes had to

¹ Otto A. L. Dieter, "Stasis," *Speech Monographs*, XVII (1950), p. 369.

say on this subject in his work *On Stases* (*περὶ στάσεων*),² written in the second century A.D.

BACKGROUND

Without going into a discussion of the various systems of stases, it should be sufficient here to remind the reader that Hermagoras is usually given credit for having written the first complete system in the second century B.C.³ Later writers on the theory of stasis fall into two general groups, those who followed Hermagoras in listing four major stases of conjecture, definition, quality, and location, and those who reduced the number of major stases to three by considering locative stasis under the heading of definition.⁴ Legal questions were sometimes included among the stases and sometimes they were discussed as questions outside the realm of stases proper.⁵ Conjectural, defini-

tive, and locative stases were regarded as being suited to questions occurring in all kinds of speaking without any necessity for special adaptation in form to those kinds; then, as now, it was recognized that the approach to agreement or disagreement on the existence of a person or thing, or on the definition of a term, is essentially the same regardless of the kind of speaking involved.

It is in the qualitative area, however, that Hermagoras and some of his successors provide a more comprehensive and separate treatment of stases of quality which are especially adapted to each of the three kinds of speaking, forensic, deliberative, and epideictic. The reason for the more thorough exposition is obvious—questions of degree are by nature more complex and, since the kinds of speaking are different in their primary aims, treatment of these more complex questions in relation to those aims is patently desirable. In ancient theory, the aim or *τέλος* of forensic oratory is justice (*τὸ δίκαιον*), the aim of deliberative oratory is expediency (*τὸ συμφέρον*), and the aim of epideictic oratory is honor (*τὸ καλόν*). It is important to note at this point, however, that the early theorists regarded the special end (or its opposite) of each type of oratory as characteristic of that type but *not as confined to it*.⁶ In a forensic speech, for instance, the special end is justice but considerations of expediency and honor may also be very much in order. Similarly, questions of justice and expediency occur in epideictic speeches, and questions

² My translation is based on the section on deliberative speaking in George Kowalski's definitive edition, *Hermogenes De statibus* (Warsaw, 1947), pp. 51-55. This passage differs in only minor detail from its counterpart in the recension of Hugo Rabe, *Hermogenis opera* (Leipzig, 1913), pp. 76-79. Hermogenes wrote a complete digest of the rhetorical theory of his time in five treatises: a set of *progymnasmata* or school exercises, and separate works on invention, stases, style, and delivery. Only those on stases and on style are generally recognized as completely the work of Hermogenes. See *Hermogenis opera*, ed. H. Rabe (Leipzig, 1913), p. iii and pp. ix-xii. According to the *Oxford Classical Dictionary*, Hermogenes exerted considerable influence, "especially in Byzantine times." A thorough study of his influence remains to be written.

³ For reconstructions of the system of Hermagoras, see C. W. Piderit, *Commentatio de Hermagore rhetore* (Hersfeld, 1839) and Georg Thiele, *Hermagoras* (Strassburg, 1893).

⁴ Among writers including locative stasis (transference) as a fourth and separate type: Hermogenes, Aurelius Augustinus, Julius Victor, Cassiodorus, Alcuin, and Clodianus. The writer of the *Ad Herennium*, Cicero (except in the *De inventione*), Apollodorus, and Quintilian turned to three stases as being sufficient.

⁵ Among writers treating legal questions separately: Apollodorus, Theodorus, Fortunatianus, Aurelius Augustinus, Julius Victor, Sulpitius Victor, Martianus Capella, Cassiodor-

us, Isidor, and Alcuin. In the *Ad Herennium*, the mature Cicero, Quintilian, and Hermogenes, these questions fall to one or more of the basic stases. For a discussion of the theory involved, see Quint. iii. 6. 86-90.

⁶ Aristotle's *Rhetoric* 1358b. See also E. M. Cope, *An Introduction to Aristotle's Rhetoric* (London, 1867), p. 120.

of justice and honor occur in deliberative speeches.

In pointing up the special importance of qualitative considerations in deliberative speeches, Cope quotes Quintilian to the effect that all the objects of deliberation are in the category of quality.⁷ In the passage to which he refers (vii. 4. 2.), Quintilian is discussing quality in a general sense, and my own interpretation of what he has to say would be less all-inclusive. However, it is certainly true that, once the relatively elementary conjectural and definitive matters have been agreed upon or waived, the bulk of a deliberative argument is normally centered on qualitative evaluations. Speakers on questions of policy infrequently contest the existence of a situation or definitive questions; they more often differ on *how great is the need for action*, on *how efficient is the plan*, and on other questions of degree.

Early writers placed a very natural emphasis on qualitative questions in all kinds of speaking, and used the kinds of speaking as convenient and logical pegs on which to hang the treatment of questions peculiar to each kind. The aim or *τέλος* of each kind was of consequent importance in determining what issues would develop. It is not surprising, then, that for each end of speaking, a standard set of subordinate headings (*κεφάλαια*) crept into common usage very early in the history of rhetorical theory and practice; these subordinate ends came to be called, for apparent reasons, the *τέλικα κεφάλαια*.⁸ They constituted a

list of topics providing an orderly treatment of any subject and offering *focal points for opposing views* on the subject. These stock topics were also stock issues.

In the area of forensic oratory, for example, a standard set of these subordinate headings provided a kind of outline for the presentation of, or attack upon, a particular plea. In the *Ad Herennium*, we are told that the subheads to be used in pleading justification (purgatio) are necessity, accident, and ignorance.⁹ It is clear that one making this plea has the option of putting forth arguments on one or more of these convenient subheads; it is equally clear that any one of these stock subheads is also a stock issue in the sense that each has potential for becoming a "center of controversy" when brought into play in a particular case. In the same way, the author of the *Ad Herennium* provides a standard pattern of subheads for guiding the inventive process as it is related to other pleas commonly used in the courts.

In the areas of deliberative and epideictic oratory, however, there were no complicated systems of intervening technical terms, and we find sets of the so-called *τέλικα κεφάλαια* developing in direct support of the primary aims of these two kinds of speaking.¹⁰ This paper is, of course, concerned only with those stock heads common to deliberative speaking. As we have seen, the principal end or *τέλος* of deliberative speaking is expediency (*τὸ συμφέρον*). What were the traditional heads or top-

⁷ Cope, *Ibid.*, p. 119.

⁸ Theodorus of Gadara (first century B.C.) is credited with popularizing the term *κεφάλαιον γενικώτατον* (most general head) as a synonym for stasis. See C. W. Piderit, *De Apollodoro . . . et Theodoro* (Marburg, 1842), p. 30. Hermogenes' use of *τέλικά κεφάλαια* (Kowalski's edition, p. 27) is usually traced to Theodorus and has been criticized as "stupid eclecticism" by Georg Thiele in his review

of H. Usener's edition of the *Ars rhetorica* of Dionysius of Halicarnassus in *Göttingische Gelehrte Anzeigen* 159 (March, 1897), p. 246. In *Pauly-Wissowa* (VIII, p. 871), Kroll denies that Hermogenes is a simple compiler and says that his choices are reasoned and defended with heat.

⁹ ii. 16. 23.

¹⁰ See Richard Volkmann, *Die Rhetorik der Griechen und Römer* (Leipzig, 1885), p. 301 ff.

ics associated with this end and serving as possible centers of controversy or stock issues?

In Aristotle's discussion of topics suited to the various kinds of speaking, he sets forth *the possible* (*τὸ δύνατον*) as a general head adapted to all forms.¹¹ He also prescribes *the expedient* (*τὸ συμφέρον*) as the specific head and aim (*τέλος*) for deliberative speaking.¹² He adds that *the just* (*τὸ δίκαιον*) and *the honorable* (*τὸ καλόν*) may also be appropriate in deliberation, although they are the aims of forensic and epideictic speaking, respectively.¹³ Thus, it is no distortion of the *Rhetoric* to say that it offers at least four topics which, as potential centers of controversy in deliberative speaking, are also convenient "stock issues."

In the less philosophical *Rhetorica ad Alexandrum*, we find a listing of the following *τέλη* to be considered in deliberative speaking: the just, the lawful, the expedient, the honorable, the pleasant, and the easy. The text adds that, if something is not easy, it may be shown to be possible; if it is not expedient, it may be shown to be necessary.¹⁴ What we have here is a list which parallels Aristotle's except that it may be said to be set up in couplets covering the same areas: (1) the just and the lawful, (2) the expedient and/or the necessary, (3) the easy and/or the possible, and (4) the honorable and the pleasant. A proposed action is measured against these standard topics or stock issues, and specific issues do or do not result, depending on whether or not a position taken on any topic is opposed.

¹¹ Aristotle's *Rhetoric*, 1391b

¹² *Ibid.*, 1358b

¹³ *Ibid.*

¹⁴ *Anaximinis Ars Rhetorica*, ed. L. Spengel (2nd ed.: Leipzig, 1850), pp. 5-6.

For centuries following the writing of Aristotle's *Rhetoric* and the *Rhetorica ad Alexandrum*, succeeding theorists have used the same patterns of deliberative topics, ends, or stock issues. The accompanying table shows how uniformly representative writers adopted essentially the same issues from the fourth century B.C. through the Renaissance. That "modern" theory and practice follow suit is quite evident from current textual treatment and from the fact that deliberative speakers still ask themselves these questions (and others related or subordinate to them) about any proposed action:

1. Is it *just*? Is it lawful?
2. Is it *expedient*? Is it useful? Is it necessary? Is it safe?
3. Is it *possible*? Is it easy?
4. Is it *honorable*? Is it pleasant?

To the stock issues above, Hermogenes (2nd century A.D.) appended that of *the anticipated effect* (*τὸ ἐκβησσόμενον*). This new addition is seldom included by other writers, except those writing commentaries on, or summaries of, earlier works.¹⁵

The writers listed in the table are *representative choices* covering a span of twenty centuries from the fourth century B.C. to the seventeenth century A.D. In all that time, no writers of consequence differed in any significant degree with those listed in their choices of stock issues for deliberative speaking; since the Renaissance, further, there have been no important departures from, or additions to, the stock issues used by, and enumerated by, the ancients. Our present-day deliberative speakers, legislators, and debaters are following a tradition that is at least 2500 years old.

¹⁵ Sopater, Planudes, *et al.*; see accompanying table.

WRITERS AND THEIR STOCK ISSUES¹⁶

| | 1a | 1b | 2a | 2b | 2c | 3a | 3b | 4a | 4b | 5 |
|----------------------|----|----|----|----|----|----|----|----|----|---|
| Aristotle | * | | * | | | * | | * | | |
| <i>Ad Alexandrum</i> | * | * | * | * | | * | * | * | * | |
| <i>Ad Herennium</i> | | | * | | * | | | | * | |
| Cicero | | | * | | | | | * | | |
| Dionysius | * | | * | | | * | | * | | |
| Longinus | * | * | * | | | * | | | | |
| Quintilian | | | * | | | * | | * | | |
| Hermogenes | * | * | * | | | * | | * | | * |
| Theon | * | | * | * | | * | * | * | * | |
| Aphthonius | * | * | * | | | * | | | | |
| Sopater | * | * | * | | | * | | * | | * |
| Syrianus | * | | * | * | * | * | * | * | * | |
| Sulp. Victor | * | * | * | | | * | | * | | |
| Cassiodorus | | | * | | | * | | * | | |
| Isidor | | | * | | | * | | * | | |
| Alcuin | | | * | | | * | | * | | |
| Emporius | * | | * | | | * | | * | | |
| Planudes | * | | * | | | * | * | | | * |
| M. Camariota | * | * | * | | | | | | | * |
| C. Soarez | | | * | | | * | | * | | |
| J. Alsted | * | * | * | * | | * | | * | * | |
| N. Caussin | | | * | * | | * | | * | * | |
| G. Vossius | | | * | * | | * | | * | * | |
| T. Farnaby | | | * | * | | * | | * | * | * |

- 1a. Is it *just* (*δίκαιον*, *iustum*)?
- 1b. Is it *lawful* (*νόμιμον*, *legitimum*)?
- 2a. Is it *expedient* or *useful* (*συμφέρον*, *χρήσιμον*, *utile*)?
- 2b. Is it *necessary* (*ἀναγκαῖον*, *necessarium*)?
- 2c. Is it *safe* (*ἀκίνδυνον*, *tutum*)?
- 3a. Is it *possible* (*δύνατον*, *possibile*)?
- 3b. Is it *easy* (*ράδιον*, *facile*)?
- 4a. Is it *honorable* (*κάλον*, *ἔνδοξον*, *honestum*)?
- 4b. Is it *pleasant* (*ἡδὺ*, *iucundum*)?
5. What is the anticipated effect (*τὸ ἐκβησόμενον*, *eventus*)?

¹⁶ Aristotle's views on topics suited to deliberative speaking are discussed in this article. In the *Ad Herennium*, *utilitas* (expediency or advantage) is the primary aim of this kind of speaking with *tuta* (security) and

honestas (honor) given as subordinate aims to which other aims are, in turn, subordinate. With the foregoing exceptions, the list of topics or stock issues shown for each author or text is the "basic list" in each case without any

As a part of that tradition, the translation which follows tells us what Hermogenes had to say about stock issues in deliberative speaking in his work on *stases*, *περὶ στάσεων*. The passage is brief but (1) it is the first extant Greek text to go into this kind of detail on the subject since initial work by Aristotle and the writer of the *Ad Alexandrum* and (2) it has never before been translated, as far as I have been able to determine, into any modern language. Hermogenes is here concerned, not with philosophical aspects of the theory of stasis, but simply with the division (*διαιρέσις*) of a deliberative question into its standard heads (*κεφάλαια*) or stock issues. The

attempt to include any further subordinate topics brought out in the discussion of those listed. In general, the topics appear in the order (from left to right) in which they were originally presented; however, I have placed the *just* (1a) in first position in the table in order to indicate its importance as a primary aim, although its conventional position in the lists is after that of the *lawful* (1b).

Excluding Hermogenes' addition of the question on *anticipated effect* (5), it is apparent that the topics in Aristotle and/or in the *Ad Alexandrum* provide the pattern for subsequent lists. It is also apparent that the major topics listed by later writers are identical with the four in Aristotle; other topics in the same areas are obviously related to those four primary aims (1a, 2a, 3a, 4a).

The asterisks in the table indicate use of a term or its equivalent by a specific writer. See citations above for Aristotle's *Rhetoric* and the *Rhetorica ad Alexandrum*. Cf. *Ad Herennium* iii. 2. ff.; Cic. *De inv.* ii. 51. ff., *De or.* ii. 82. 334. ff.; Dionysius, *Ars rhetorica*, ed. H. Usener (Leipzig, 1895), p. 119 f.; Longinus fr. 15; Quintilian iii. 8. 22. ff.; Hermogenes, translation accompanying this article; Theon, in *Rhetores Graeci*, ed. C. Walz (Stuttgart, 1832-36) I, 244; Aphthonius, Walz I, 115; Sopater, Walz V, 182; Syrianus, Walz IV, 701; Sulp. Victor, in *Rhetores Latini minores*, ed. C. Halm (Leipzig, 1863), p. 342; Cassiodorus, Halm p. 501; Isidor, Halm p. 508; Alcuin, Halm p. 527; Emporius, Halm p. 571; Planudes, Walz V, 336; M. Camariota, Walz VI, 624; Cyprianus Soarez, *De arte rhetorica* (Cologne, 1570), p. 62; Johann Alsted, *Orator* (Herborn, 1612), p. 96 f.; Nicolas Caussin, *De eloquentia* (Cologne, 1681), p. 813 f.; Gerardus Vossius, *Commentariorum rhetoriconum libri vi* (Leyden, 1643), p. 32 f.; Thomas Farnaby, *Index rhetoricus* (London, 1633), p. 7 f.

reader will readily see the parallels with "modern" techniques which I have indicated in the foregoing introduction.

TRANSLATION

On A Deliberative Question

The deliberative question *ἡ πραγματική*,¹⁷ is divided into (these heads or stock issues): the lawful (*νόμιμον*), the just (*δίκαιον*), the expedient (*συμφέρον*), the possible (*δύνατον*), the honorable (*ένδοξον*), and the anticipated effect (*έκβησόμενον*).

The deliberative question can be written or unwritten. The written is that in which the question is derived from the letter of a document; for instance, a law required deliberating about war for three days but, when Philip occupies Elatea, Demosthenes immediately enters a motion for marching out. The unwritten is that in which the question is not derived from the letter; for instance, after (his victory at) Pylos, Cleon asks to be called Pythius.

Within the written classification, to continue, (the stock issue of) the lawful falls under one of the legal issues and will be divided according to (the standard pattern for) it; we shall soon take up the legal issues.¹⁸ Within the unwritten classification, general custom must be examined as being akin to law; for instance, (in answer to Cleon) "You ask for unprecedented things both in such kind and in such quantity as no one ever before." We shall treat these things, in turn, from the standpoints of direct-opposition (*ἐνστάσις*) and of in-

¹⁷ Greek terms included in the translation and in the notes appear in the nominative case for easier comparison with the same terms appearing in the preceding section on background.

¹⁸ Hermogenes deals with the usual and traditional legal issues of letter and intent, conflict of laws, inference, and ambiguity. See Kowalski, pp. 58-70.

direct-opposition (*ἀντίπαραστάσις*). (Here is a reply) in indirect-opposition to "You ask for unprecedented things," since Cleon is asking to be called Pythius after those events around Pylos: "I (Cleon) do ask for unprecedented things, for the things done are also unprecedented." (Here is a reply) in direct-opposition: "If it comes to that, these things are not at all unprecedented, for Pericles was called Olympius." Which of the two must be used first, direct-opposition or indirect-opposition, will be dictated by the circumstances, as we also directed in connection with the conjectural issue.¹⁹

(The stock issue of) *the just* falls under one of the forensic issues and, so far as it is possible, it will be divided according to the heads of that classification.²⁰

(The stock issue of) *the expedient* is dual in reference to (considerations of what is) useful (*χρήσιμον*) and necessary (*ἀναγκαῖον*); for instance, "It is useful to accept Olynthus as an ally but it is even more necessary, lest Philip become powerful against us through his overcoming them (the Olynthians)."²¹ You will examine this issue in two ways —(1) what will follow our doing any particular thing about which there is deliberation and (2) what will follow our not doing so. Then, you will ex-

¹⁹ Kowalski, pp. 22-23. In this passage, Hermogenes discusses the use of direct-opposition and indirect-opposition as subdivisions of objection (*μετάληψις*) to a plea of confession-and-avoidance (*ἀντίληψις*).

²⁰ Kowalski, pp. 40-51. Like earlier writers on the theory of stases, Hermogenes divides the forensic issues into those of (1) confession-and-avoidance and those of (2) antithetical countercharges (*ἀντίθετικα*).

²¹ A paraphrase of the general theme pervading the *Olynthiacs* i, ii, and iii. (So far as I can discover, all quotations in this section of Hermogenes are paraphrases.)

amine each of these alternatives in turn in four ways. For example:

"If we should take it (Olynthus) to ourselves, (a) those desirable things existing for us will endure, and (b) those desirable things not in being will be ours; in turn, (c) we shall get rid of the existing undesirable things, and (d) we shall not take to ourselves those undesirable things not in being.

"But if we should not take it to ourselves, (a) those desirable things existing will be destroyed, and (b) those desirable things not in being but about to be acquired will not be ours; in turn, (c) those existing undesirable things will endure, and (d) those undesirable things not in being will be ours."²²

To go on, you will subdivide (the stock issue of) *the possible*, first by showing through the use of direct-opposition that it (proposed action) is not difficult (*οὐ χαλεπόν*), and, then, through the use of indirect-opposition that, even if it is difficult, it is nevertheless necessary; further, that there is need to endure both sufferings and dangers on behalf of worthwhile things, and thereby, avoid falling into worse evils; then, that it (proposed action) is really easy (*ράδιον*). However, if it (the deliberation) is about war or something of this kind, examine *the possible* from the standpoint of the attributes of the persons (involved), just as Demosthenes did in the second *Olynthiac*—how Philip feels, that he is disheartened, or in what condition are his external affairs, such as Thessalians, Illyrians, finances, "the household troops and footguards about him," and so on.²³ The careful review of the power of the Athenians was (i.e. could have been) antithetical to the foregoing, but one examining this speech from the standpoint of theory will readily bring to light why the ora-

²² Both paragraphs are again paraphrases of the arguments presented in the three *Olynthiacs*.

²³ Cf. *Olynthiacs* ii. 14-21.

tor omitted it;²⁴ however, the division does (technically) call for it. We should not be surprised if it is passed over, since some of the other headings were also quite naturally omitted there, and we ourselves often leave them out, not only in deliberations but elsewhere as well; for the system shows the method of division, but one must apply it to the nature of the thing done in accordance with what is permissible—for the one has to do with invention, but the other has to do with judgment and, further, we do not say whatever we discover but whatever we choose to be spoken.

(The stock issue of) *the honorable* (is treated) in a manner similar to that used for *the expedient*; for instance:

"(a) What glory now existing will remain for us, (b) what glory not in being will result if we should do these things about which there is deliberation; (c) what ignominy in being shall we be thrusting aside, and (d) what ignominy not in being but expected shall we be avoiding?"

²⁴ Early in the *Second Olynthiac*, Demosthenes says he does not choose to list the resources of Philip because "such a topic is a . . . record of Athenian failure." Since he does not go into detail on the resources of Philip, presumably there is no need to review Athenian power as antithetical to them. See *Olynthiacs* ii. 3.

"On the contrary, if we should not do these things, (a) what disgrace in being shall we not put aside, (b) what disgrace not now in being will accrue to us, (c) what glory existing for us will disappear, and (d) what glory not existing but expected will not come to pass?"

(The stock issue of) *the anticipated effect* consists of speculation by hypothesis in regard to each of the things coming to pass as a result (of the proposed action); for example, "Whether or not we shall prevail, it is better to have given aid but, for the present, there is profit in only having voted (upon it)."²⁵ Again in the *Philippics*, for instance, "It (the voting) will most certainly not be a thing to be scorned, by Zeus, even if you will not do as I myself recommend, so that, aware that you are prepared, he may stay away through fear or, disregarding these things, he may be left off his guard."²⁶ And again, "without us, whether the allies needing our aid will be conquered or whether they will conquer, they will be harmful in either case," as in the speech on behalf of the Megalopolitans.²⁷

²⁵ Cf. *Philippics* i. 18.

²⁶ See the *Philippics*, especially the conclusion to *Philippics* iii.

²⁷ Cf. *For The People of Megalopolis* 30.

USE OF AUTHORITIES AS ETHICAL PROOF IN THE TALMUDIC DISCOURSE*

GERALD M. PHILLIPS
North Dakota State College

AUTHORITIES IN LOGICAL PROOF

ARISTOTLE, in speaking of *ethos*, states in Book One, Chapter Two of the *Rhetoric*, "The first kind (of persuasion) depends on the personal character of the speaker." He goes on to elaborate:

[Persuasion is] achieved by the speaker's personal character when the speech is so spoken as to make us think him credible. We believe good men more fully and more readily than others. This is true, generally, whatever the question is, and absolutely true where exact certainty is impossible and opinions are divided. This kind of persuasion, like the others, should be achieved by what the speaker says, not by what people think of his character before he begins to speak. It is not true, as some writers assume in their treatises on rhetoric, that the personal goodness revealed by the speaker, contributes nothing to his power of persuasion; on the contrary, his character may almost be called the most effective means of persuasion he possesses.

It appears, from this statement by Aristotle, that according to the classical Greek concept, ethical proof was regarded as probably the most effective form of proof. That while it was held desirable that audiences make their judgments based on logical presentation of the facts of the case, in practice, the character of the speaker or the impression that the audience has of the speaker as a man appeared to be the most important and most effective means of persuasion.

Though the rhetorical theory of the talmudic rabbis was not explicitly ex-

pressed, and while there is no clear-cut delineation of their mode or theory of rhetoric as there is in the Greek culture, through study of their practice of rhetoric it is possible to ascertain that ethical proof occupied as high a rank in the talmudic discourse as it apparently occupied in the Greek discourse. According to Professor George Foot Moore, the task of the talmudic rabbi was "to discover, elucidate, and apply what God teaches."¹ The use of the talmudic lecture as a means of communicating "law"² involved the rabbis in proving conclusions to their students, and, like their Greek counterparts, they used both inductive and deductive forms of proof. In essence the forms of proof used by the rabbis were based on (1) observed fact, (2) testimony of witnesses, (3) statements from recognized authorities or authoritative works, and (4) analogies drawn from similar circumstances, similar sentences, phrases, words, letters, and marks in the Biblical text. Many of the forms of logical proof they employed were highly artificial and bore no relationship to the formal logical systems set up and employed by the Greeks.³ However, in the use of witnesses and authorities there are a

¹ George Foot Moore, *Judaism* (Cambridge, 1944), Vol. I, p. 248.

² "Law" in this case, means the application of Divine revelation to specific instances. This application is known as *halachah*, from the Hebrew verb "halach," meaning literally, "a way to go."

³ E. g. proofs seeking to draw analogies through occurrence of similar words and phrases in dissimilar Bible passages. Frequently this resulted in a highly strained sort of proof.

*Based on Ph.D. dissertation, Western Reserve University, 1956, directed by Robert A. Lang.

reasonable number of similarities between Greek and Hebrew usage, and certainly, these two kinds of logical proof appear to be the most frequently used and also the most important forms of proof employed by the talmudic rabbis.

Most of the deductive proof in talmudic oratory had as its basis either Scripture or authoritative material relating to Scripture. This material included: (1) *Baraita*, a collection of relatively unimportant oral laws left out of the official code; (2) *Midrash*, a collection of past synagogue sermons; (3) *Mishnah*, the authoritative Oral Law,⁴ and (4) the statements of past and present authorities. The citing of a conclusion drawn from the general statement of a valid authority appeared to be the form of proof most frequently used by the rabbis. It was considered essential, however, that the reasoning process of the authority be presented along with his general statement in all cases where the authority did not quite qualify as an unimpeachable source. The simple use of the name of an authority in many cases was not considered sufficient proof for a conclusion. One of the basic ethical maxims of the Talmud states:

Rabban Gamaliel used to say: Appoint a teacher for thyself and avoid doubt (select an authority who is respected; who can be consulted in cases of difficulty) and make not a habit of tithing by guesswork (know the facts and reasons for a decision before giving a decision).⁵

⁴ This was the "Oral Law" handed down by the pre-Christian era Scribes. Rabbinic tradition attributes this "Law" to Moses, who did not cause the entire "Law" to be written down when it was revealed at Mt. Sinai. Rabbinic belief is that, though *Mishnah* is related to and dependent upon Scripture, it is equally authoritative.

⁵ *Abot I, 13.* Citations to Talmud hereafter include the reference to the standard Hebrew-Aramaic editions (the first number) plus a reference to English translation: *The Babylonian Talmud*. Dr. I. Epstein, general editor (35 Vois.) London: The Soncino Press, 1936-1952.

This statement indicates that although the rabbis held their authorities in high regard, they considered it necessary, in most cases, to present the reasons for the conclusion drawn by the authorities. In practice, however, many authoritative sources like the Bible, *Mishnah*, and certain individual rabbis were accepted at face value by their listeners without requiring further proof.

Scripture was considered the highest authority. Even when a recognized scholar presented an opinion, it was necessarily related to Scripture so that the listener could recognize the authoritative religious basis for the statement.⁶ Even *Mishnah*, though equally authoritative, still was related to Scripture, either as an interpretation or as an explanation.

Since the purpose of the lecture in the Talmudic Academy was to apply Scripture to every-day situations to determine *halachah* (the application of scriptural law to any given case), it was necessary that logical proof for the conclusion as to what was *halachah* have some relation to Scripture. Many rabbis considered a simple scriptural exposition made by a recognized authority as a superior, irrefutable proof⁷ and, in fact, one authority, R. Akiba, once said, "If this is an authentic tradition, we shall accept it; but if it is only a logical deduction, there is a rebuttal."⁸ The implication here is that the association of the name of a recognized authority together with his scriptural proof makes a statement irrefutable. Since Scripture was held to be divine, statements relating to Scripture, if made by humans who were recognized as authorities, were also held to be divine. It was from this process of elevation of human

⁶ *Kethuboth* 83b, 528.

⁷ *Zebahim* 48a, 239.

⁸ *Kerithoth* 15b, 118.

authorities to approach divine authority that the rabbinic concept of ethical proof emerged.

The use of authorities commencing with a form of logical proof and culminating in a form of ethical proof resulted in the establishment of a clear-cut hierarchy of authority which determined the exact degree of authority a given individual had. For example, the Scribes (the early authors of the Oral Law) could not dispute the Scripture; a *Tanna* (later interpreters of the Oral Law) could not dispute the direct word of a Scribe; and an *Amora* (responsible for the commentary on the Oral Law or *Gemara*) could not reverse a direct statement of a *Tanna*, and within each category certain individuals outranked others.⁹ The ranking of authorities was determined by a set of generally accepted criteria: (1) determination of whether the authority was a teacher or a student since the authority of a teacher was superior in all cases,¹⁰ and (2) the size and reputation of the school with which the authority was associated.¹¹ Further tests applied to an authority were his familiarity or lack of familiarity with the material under discussion and the field in which he was held to be an expert. R. Ashi, for example, was considered the final talmudic authority since he was the head of the last academy and was popularly believed to be the redactor of the Talmud.¹² R. Eliezer, who only quoted the opinions of others, was considered a highly valid authority since in the few cases where he did render a decision, his views were regarded as "thoroughly sifted" and where his opinion clashes

with that of another authority, his opinion is regarded as correct.¹³ R. Zera cautioned his scholars to reject any statement of *Baraita* that was not traceable directly back to R. Hiyya and R. Oshaiah, since these were held to be the men who made the original collection of *Baraita*, and were, therefore, the recognized authorities in that field.¹⁴ Certain individuals were regarded as poor authorities because of misstatements or errors previously made. For example, R. Abiathar was considered to be a poor authority since he was frequently caught in misstatements of facts.¹⁵ Rabbah b. Bar Hana was regarded as the champion liar of the Talmud, and his opinions were invalid in all matters; but he was regarded as worth listening to since his tall tales were interesting.¹⁶

A further qualification of authorities was made on the basis of the personal reputation of the man. It is at this point, too, that a closer approach to a concept of ethical proof is made. For example, R. Hiyya was regarded as "accurate in reporting the statements of his teacher."¹⁷ Accuracy in reporting was held to be a highly desirable personal quality. R. Eleazar was known as "a great man" and, therefore, as a worthy authority.¹⁸ R. Johanan b. Zakkai was an accepted authority because "during his whole life he never uttered profane talk nor walked four cubits without studying the *Torah*."¹⁹ R. Zera pointed out that a valid authority frequently engaged in reviews with his teacher and was constantly checking the validity of his opinions.²⁰ It was generally accept-

⁹ *Yadayim* III, 2.

¹⁰ *Niddah* 7b, 45.

¹¹ *Rosh Hashanah* 31b, 151.

¹² Zevi Hirsch Chajes, *The Student's Guide Through The Talmud*. Translated by Jacob Schacter. London: East and West Library, 1952, p. 79.

¹³ *Kethuboth* 40a, 223.

¹⁴ *Hullin* 141a, 818.

¹⁵ *Gittin* 6b, 20.

¹⁶ *Yebamoth* 55b, 373.

¹⁷ *Berakoth* 33b, 208.

¹⁸ *Kethuboth* 53a, 312.

¹⁹ *Sukkah* 28a, 122. "Torah" means both the Pentateuch and the whole body of Jewish law.

²⁰ *Berakoth* 36b, 240.

ed that the opinion of the head of an academy would have more weight than that of any of his associates while the direct statement of an authority had more weight than a reported statement of an authority.²¹ Through these individual statements qualifying authorities, a partial view of what was held to be desirable in an individual could be ascertained; and it is from these individual statements that the rabbis eventually evolved for themselves the concept of an ideal authority. The ideal rabbi was physically large, had prestige, rank, family connection, knowledge; accuracy, self-criticism, and clean talk were characteristic.

If an individual qualified as an authority acceptable for proof, it was usually mandatory that his views be accepted, at least by his students and disciples. He had, of course, the obligation of presenting some sort of proof, but really it was immaterial in practice what particular proof was presented since there are frequent examples given in the Talmud of alternate proofs that an authority could have presented in order to arrive at the same conclusion.²² As a matter of fact, one of the standard exercises at the academies consisted of having the students devise alternate proofs for the same conclusion. It is important to note, however, that the conclusion originally stated by the authority must always result from the alternate proofs. Thus, conclusions drawn from the statements of an authority were superior to conclusions drawn through pure reason not based on any recognized authority. No matter how sound the reason, individual argument could be refuted by citing the documented opinion of a recognized authority.²³

²¹ Niddah 14b, 95.

²² Sanhedrin 36a, 228.

²³ Zebahim 96b, 461.

If a speaker presented no valid authority for his view, he was then obligated to present logical proof. If no authority could be found by those on the opposite side, then the question could be debated on its logical merit. However, if an authority could be found on the opposite side, then the original view was held to be refuted.²⁴

The talmudic rabbis in all of their disputations used majority rule to resolve conflict. Their doctrines of majority rule also applied to the use of authority in proof. For example, when several authoritative statements could be found which agreed on a given *halachah*, then these statements determined the *halachah*.²⁵ Throughout the Talmud, when the opinion of an individual was presented along with an anonymous opinion, the anonymous opinion was held to represent the view of the majority of the authorities, and was, therefore, considered to be the valid, accepted view.²⁶ Though the opinions of individuals in conflict with the majority were sometimes cited, they were cited in accordance with a doctrine of minority freedom and are not to be considered as sources of valid law.²⁷

The rabbis accorded a great deal of respect to the opinions of living authorities. In any practical debate, the opinion of one authority could not reverse the opinion of a contemporary authority who had ruled on the question previously. This was based on the traditional theory that no assembly or individual could annul the decisions of another assembly or individual unless it was superior in both wisdom and number. If two equal authorities clashed, they were presumed to be equal in wis-

²⁴ Sanhedrin 90b, 187.

²⁵ Yebamoth 40b, 276; Niddah 49a, 340.

²⁶ Yebamoth 40b, 276.

²⁷ Edduyoth I, 5.

dom, and since neither was superior in numbers, the view of one could not over-rule the view of another.²⁸ When two contemporary authorities clashed on this basis, the process of majority vote by the associates of the academy was used to determine the law; and if no vote was taken, the problem was held to be unsolved and left for a later generation.²⁹ In these conflicts of contemporary authorities a form of seniority evolved, so that the authority who had made the first ruling on a given item of law was held to be the authority whose ruling stood until reversed by an assembly that was superior in numbers or in wisdom. It is somewhat difficult to ascertain what "superior in wisdom" meant.

USE OF AUTHORITIES IN TALMUDIC DISPUTATIONS

The actual use made of authorities as a form of proof appeared to conform to a generally accepted methodology. Phrases like the following are characteristic of the method used by the rabbis to cite authorities in discourse:

R. Simeon b. Zebid said in the name of R. Isaac b. Tabla in the name of R. Hiyya Areka of the school of R. Aha in the name of R. Zera in the name of R. Eleazar in the name of R. Hanania in the name of R. Mi'asha on the authority of R. Judah b. Ilai³⁰

The rabbis took pains to report the sources of their authoritative statements accurately. The long list of names in the example indicates that the opinion presented was a view of long standing, and the names cited trace it back to its origin, while the name of each authority who previously cited it gives the opinion added weight.³¹ The rabbis were required to repeat statements of

authorities as they were made. If they did not hear the statement directly from the authority, they had to cite the names of the persons from whom they heard it so that if an inaccuracy was discovered, it could be traced back to its source.³² It was not considered sufficient to capture the "spirit" of the statement of the authority for "a man must cite . . . in the exact language of his master."³³ There is a case recorded in the Talmud where a lecturer made an error in pronunciation which was reported in his name later exactly as he made it. The quoted error, based on his authoritative statement of an recognized authority, then became acceptable as a basis for later proof.³⁴

The peculiar institution of the *meturgeman* in talmudic debate was partly responsible for the stress on exact citation of sources. Because of the fact that many of the students at the academies were not fluent in Hebrew, many of the rabbis lectured through a *meturgeman*. The rabbi stated the heads of the discourse in Hebrew and the *meturgeman* translated into the vernacular and elaborated upon it. In order that the students recognize that the statements were not original with the *meturgeman*, he was required to cite the sources of his remarks with great care.³⁵

R. Eliezer was one of the most highly regarded talmudic authorities and his claim to fame was that he was entirely unoriginal. R. Eliezer boasted that he never said a word in presenting a proof unless it was an exact quote from a previous authority, and that he never held an opinion unless another authority had held the opinion before him.³⁶

²⁸ Hullin 44b, 238.

²⁹ Megillah 2a, 2.

³⁰ Nedarim 8b, 19.

³¹ Abodah Zarah 16b, 83.

³² Hullin 7a, 27.

³³ Bekoroth 5a, 23.

³⁴ Shabbath 15a, 61.

³⁵ Kiddushin 31b, 154.

³⁶ Sukkah 27b, 121.

The rabbis were quite jealous of their own statements. They expressed great concern when other persons reported their statements and did not give them due credit for them. There appeared to be a generally accepted method of oral footnoting of authority, and in cases where these references were not given, severe criticism was leveled against the speaker. Two notorious critics, R. Ammi and R. Assi, attended lectures in order to catch errors in reporting authorities.³⁷ On the positive side the rabbis were praised when they discovered the author of an anonymous statement.³⁸ When two conflicting statements about the same topic were reported in the name of one authority, neither statement could be accepted until the correct and original statement by the authority had been discovered or ascertained.³⁹

In order to follow the clash of authorities in disputation, identifications are given in the Talmud. Certain individuals who were highly regarded as authorities were not cited by name but were cited by euphemism. For example, the phrase "our rabbis in Babylon" referred to Rab and Samuel. The phrase "the keen intellect of *Pumbebitha*" referred to Efa and Abimi. The phrase "they taught it in the West" referred to R. Eleazar.⁴⁰

AUTHORITIES IN ETHICAL PROOF

There are several citations *passim* the Talmud qualifying certain authorities as superior in every case of clash with another authority. Although there is no one place where this entire table of authority is mentioned throughout the Talmud, statements occur which

³⁷ Yoma 73a, 350.

³⁸ Sukkah 11a, 43.

³⁹ Kethuboth 27a, 336.

⁴⁰ Sanhedrin 17b, 88. V. also Horayoth 13b, 102 for further identification of authorities.

qualify one authority above another. For example, the opinion of R. Johanan b. Zakkai takes precedence in every case over the opinion of Rab and Samuel.⁴¹ It is also stated that the opinion of R. Judah Ha-Nasi takes precedence over the opinion of any of his contemporaries or followers.⁴²

Most authorities were granted superiority over specific other authorities in case of clash. There was, however, one rabbi who was held to be the authority *par excellence*. This was Hillel who provides the model for the whole scale of authorities. Hillel is held to be the prototype of the ideal rabbi, and hence, the ideal speaker. He was placed in juxtaposition with Shammai who reveals all the opposite characteristics; and so, by this contrast, Hillel's good qualities and Shammai's bad qualities tend to establish the ideal of rabbinic authority. Virtually, every talmudic debate brings in some matter of difference between Hillel and Shammai.⁴³ In each of these clashes, Hillel was victorious, and the rule was established that the view of Hillel was supreme in all cases.⁴⁴ This view of Hillel's authority was further reinforced by an alleged declaration from heaven "that the law is with Hillel."⁴⁵ The elevation of Hillel to the highest authority, provides the precedent for the later elevation of other rabbis to authoritative roles.

The Talmudic concept of the use of authority as a superior form of proof, derived its sanction from one of the legends about Hillel. According to

⁴¹ Bezaḥ 5a, 17.

⁴² Sanhedrin 36a, 337. For further identification of rank of authorities *v.* Berakoth 26b, 161; Kethuboth 100a, 634; Yadayim IV, 3; Kiddushin 39a, 189; Bezaḥ 27a, 137; Erubin 46b, 323; 47b, 327; Shabbath 21b, 91; Hullin 77a, 427.

⁴³ Berakoth 51b, 311.

⁴⁴ Yebamoth 9a, 43. Hillel's view was superior in all but four specific instances discussed *infra*.

⁴⁵ Erubin 13b, 85.

this legend, when Hillel first arrived in Jerusalem, the Elders were engaged in a dispute about a difficult point of law. Hillel settled the dispute by use of logic. The Elders demanded he present an authority to support his logical proof and Hillel cited the scriptural chapter and verse to support his view, pointing out that the Elders were justified in asking for this authority since questions of such importance must be settled by proof and authority, not only by one of them.⁴⁶ It is a frequent occurrence then, in the Talmud, for a rabbi who has just presented a logical proof to be asked for a concurring authority, and the rabbi who presents authority alone is generally asked to cite the proof. The principle appears to be that authority without proof is sometimes weak, while proof without authority is always weak.⁴⁷ Many of the rabbis felt that statements from great authorities of the past coupled with their original proofs were irrefutable, but any logical deduction without an authority allowed for a rebuttal.⁴⁸ The scholar who based his decisions on reports of accepted authorities is held to be superior to the original thinker or the quoter of opinions.⁴⁹ Rabina pointed to himself as typifying what is desirable in an authority by saying, "I am neither a self-pretended scholar (speaking on my own authority) nor a visionary (telling stories) nor unique, but I am a teacher (who reports correctly the statements of others) and a systematizer of traditions (who provides reasons for the views of the authorities)."⁵⁰

The extensive use of, and high regard for authorities in logical proof leads to

the talmudic concept of ethical proof. While there was no concept of the "good" orator, the prestige of individual authorities leads to an understanding of what was considered personally desirable in a speaker. The head of any academy held immense prestige in his community. He was regarded as a leader of the community while alive and as a sage when dead. Frequently, because of his prestige, his views and opinions went unchallenged despite the existence of the doctrine of logical proof that the statement of an authority must be accompanied by proof. The views which he presented were regarded as valid because it was understood that as head of an academy he was a learned man; and because of this, it was presumed that any conclusions he drew would be based on correct logic. Although many heads of academies were meticulous about their observance of the unwritten law calling for the presentation of proof for their conclusion in order to set an example for their students, some of them succeeded in having their opinions accepted simply on the strength of their position as head of the academy.⁵¹ The implication appears to be that, at least as far as the talmudic rabbis were concerned, there were certain qualities in a man which would make his word acceptable at all times. There is no statement as to the personal qualifications of a good orator, but there are numerous references to the qualities desirable in a rabbinic authority. All of these statements appeared to be related in some way to learning and scholarship. Since the rabbis displayed their learning in the academy, it appeared that the rabbis acceptance as a lecturer was based to a large extent upon his acceptance as a scholar.⁵²

⁴⁶ Pesahim 66a, 333.

⁴⁷ Erubin 46a, 318.

⁴⁸ Kerithoth 15b, 118.

⁴⁹ Horayoth 13b, 102; Gittin 67a, 317.

⁵⁰ Pesahim 105a, 550.

⁵¹ Erubin 41a, 282.

⁵² Erubin 54a, 377.

One authority points out that a rabbi must have good breeding, but without knowledge of *Torah*, good breeding is impossible; that a good rabbi must have fear of God, but without wisdom, fear of God is impossible; that a good rabbi must have understanding, but without knowledge, understanding is impossible.⁵³ These are typical of the many statements about the attributes of a rabbi. While many of the authorities disagree on the minor attributes of a good rabbi, they all appear to agree that the essential attributes are learning, scholarship, knowledge. R. Johanan, who considered himself to be the most handsome of rabbis, pointed out that a rabbi must have knowledge and a good appearance.⁵⁴ Several authorities who were descended from good families felt that a rabbi must be learned and come from a good family.⁵⁵ A large group of rabbis set up *Nahum* of *Gimzu*, who was noted for the fact that he was horribly ugly and a quadriplegic, as their ideal of a rabbi.⁵⁶ They held that it was his ugliness that forced him to acquire his great learning.⁵⁷ Most authorities agreed that both meekness and pride were desirable in a rabbi, but these could not be achieved without learning.⁵⁸

There was some variation in the views on what was to be included in the learning that a rabbi must possess. R. Dimi stated that a rabbi must be a master of Bible and *Mishnah*; he must be an expert in civil, criminal, and religious law and be able to pass judgment in "strictest accord with the truth."⁵⁹ Raba felt that learning meant the ability to understand the subject matter of Bible and *Mishnah* plus the ability

to teach it to others.⁶⁰ Some authorities held that a rabbi must have memorized the whole *Mishnah* and be able to draw conclusions from it through the use of "keen dialectics."⁶¹ There was general agreement that if a rabbi used his knowledge for personal gain, then he was not really wise.⁶²

The establishment of Hillel as the leading authority in logical proof also set him up as an ethical model to which other authorities were required to conform. Virtually, every other authority was compared to him in some way. Hillel displayed the characteristics of mastery of subject matter, great wisdom in judgment, humility, and patience. As has been pointed out above, a voice from heaven declared that Hillel was always right.⁶³ However, in usual talmudic fashion, at another point care is taken to indicate that Hillel's counterpart and great opponent, Shammai, was right in four cases to indicate human fallibility and the impossibility of one man always being right.⁶⁴ A story is told about Hillel which illustrates the ideal traits of the rabbi.

Our rabbis taught: A man should always be gentle like Hillel not impatient like Shammai. It once happened that two men made a wager with each other saying, "He who goes and makes Hillel angry, shall receive 400 zuz." Said one, "I will go and insult him." That day was the Sabbath eve and Hillel was washing his head. He went, passed by the door of his house and called out, "Is Hillel here? Is Hillel here?" Thereupon he robed and went out to him saying, "My son, what do you require?" "I have a question to ask," said he. "Ask, my son," he prompted. Thereupon he asked, "Why are the heads of Babylonians round?" (This was a gross insult since Hillel was a Babylonian and further since he addressed Hillel without a title and disturbed him while he was preparing

⁵³ *Aboth* III, 17.

⁵⁴ *Baba Mezi'a* 84a, 479.

⁵⁵ *Hullin* 92a, 514.

⁵⁶ *Ta'anith* 21a, 104.

⁵⁷ *Ta'anith* 7a, 26.

⁵⁸ *Yoma* 22b, 102.

⁵⁹ *Hagigah* 14a, 84.

⁶⁰ *Shabbath* 63a, 296.

⁶¹ *Erubin* 67a, 466.

⁶² *Baba Mezi'a* 23b, 28.

⁶³ *Berakoth* 11a, 62; *Yabamoth* 9a, 43.

⁶⁴ *Erubin* 13b, 85.

for the Sabbath.) "My son, you have asked a great question," replied he. "Because they have no skillful midwives." He departed, tarried awhile, and returned and called out, "Is Hillel here? Is Hillel here?" He robed and went out to him saying, "My son, what do you require?" "I have a question to ask," said he. "Ask, my son," he prompted. Thereupon he asked, "Why are the eyes of the Palmyreans bleared?" "My son, you have asked a great question," replied he. "Because they live in sandy places." He departed, tarried awhile, returned, and called out, "Is Hillel here? Is Hillel here?" He robed and went out to him saying, "My son, what do you require?" "I have a question to ask," said he. "Ask, my son," he prompted. He asked, "Why are the feet of the African wide?" "My son, you have asked a great question," said he. "Because they live in watery marshes." "I have many questions to ask," said he, "but fear that you may become angry." Thereupon he robed and sat before him and said, "Ask all the questions you have to ask." "Are you the Hillel who is called *Nasi* (prince) of Israel?" "Yes," replied he. "If that is so," he retorted, "may there not be many like you in Israel?" "Why, my son?" queried he. "Because I have lost 400 zuz through you," complained he. "Be careful of your moods," he answered. "Hillel is worth it that you should lose 400 zuz and yet another 400 zuz through him yet Hillel shall not lose his temper."⁶⁵

⁶⁵ Shabbath 30b, 138.

The qualities shown by Hillel in this legend were the qualities that the other rabbis tried to emulate. His teaching methods were assiduously followed and they represented the model for all talmudic teaching.⁶⁶ Hillel's ethics became the core of talmudic ethics.⁶⁷

It is apparent that the talmudic rabbis relied heavily on authorities in proof. Since the use of authority in logical proof resulted in a form of ethical proof, it appears that the talmudic rabbi accorded a rank to it, similar to that given it by the Greek rhetorician.

Despite the fact that the rabbis had no formal theory of rhetoric, and were probably unaware of the existence of the Greek concept, the development of the concept in practice appeared to follow the Greek model. The rabbis succeeded in building their concept of authority into a sort of science, so that it represents the most important single element in talmudic oratory.

⁶⁶ Shabbath 31a, 139.

⁶⁷ Aboth I, 13 for examples of Hillel's ethical maxims.